

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2005, 14:42:03 ; Search time 70 Seconds
(without alignments)
143.654 Million cell updates/sec

Title: US-10-714-852-1

Sequence: 1 NAKATYEALKQTEADLAAYKXANAA 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: geneseqp19808:*\n2: geneseqp19908:*\n3: geneseqp20008:*\n4: geneseqp20018:*\n5: geneseqp20028:*\n6: geneseqp20038:*\n7: geneseqp20038:*\n8: geneseqp20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	100.0	26	8	ADM32815
2	124	100.0	26	8	AD034199
3	124	100.0	1561	2	AAW02098
4	124	100.0	1561	6	ABU44355
5	77	62.1	1566	2	AA79643
6	64	51.6	19	1	AA80522
7	57	46.0	1286	4	ABR70681
8	57	46.0	1286	4	AD000993
9	56	45.2	16	5	AAU93845
10	56	45.2	16	7	AD010835
11	56	45.2	16	7	ADM39740
12	56	45.2	16	8	ADG63935
13	56	45.2	16	8	ADP73469
14	56	45.2	16	8	ADRI2762
15	56	45.2	17	2	AAV29700
16	56	45.2	17	5	AAU93863
17	56	45.2	17	7	AD010939
18	56	45.2	17	7	AD010939
19	56	45.2	17	8	ADG54010
20	56	45.2	17	8	ADP73580
21	56	45.2	17	8	ADRI2866
22	55	44.4	1361	8	ADN20556
23	54	43.5	982	2	AAV55955
24	54	43.5	982	8	AD031883
25	54	43.5	982	8	ADQ26204

ALIGNMENTS

26	53	42.7	248	7	ADH87955	Adh87955	Enterococ
27	52	41.9	1136	5	AB548878	Ab548878	Listeria
28	52	41.9	1136	6	ABU32894	Abu32894	Protein e
29	52	41.9	4010	4	AB861520	Ab861520	Drosophila
30	51	41.1	496	6	ABU19616	Abu19616	Protein e
31	51	41.1	1637	5	AB879896	Ab879896	Human kin
32	51	41.1	1637	5	AAU97891	AAU97891	Human kin
33	51	41.1	1637	6	AB872055	Ab872055	Human kin
34	51	41.1	1638	6	ABR39807	ABR39807	Human SCA
35	51	41.1	1638	7	ADC10130	ADC10130	Human NOV
36	50	40.3	1318	6	AA384418	AA384418	Arabidops
37	49.5	39.9	85	8	ADG31270	ADG31270	Novel mou
38	49.5	39.9	138	5	AA680146	AA680146	CYP40 dom
39	49.5	39.9	370	2	AA734196	AA734196	Mouse cyc
40	49.5	39.9	370	2	AD859004	AD859004	Human Pro
41	49.5	39.9	370	8	ADG31269	ADG31269	Novel mou
42	49.5	39.9	370	8	ADH09500	ADH09500	Human hos
43	49.5	39.9	370	8	ABM79754	ABM79754	Expressio
44	49.5	39.9	866	6	ABU18259	ABU18259	Protein e
45	49	39.5	96	6	ADA34796	ADA34796	Actinobac

RESULT 1
ADM32815 standard; peptide; 26 AA.

ADM32815;

03-JUN-2004 (first entry)

Synthetic peptide antigen derived from protein antigen serotype C.

protein antigen serotype C; PAC; surface protein antigen; carries;

antigen; secretory immunoglobulin A; saliva.

Streptococcus mutans.

Synthetic.

EP1400809-AL.

24-MAR-2004.

15-SEP-2003; 2003EP-00020492.

19-SEP-2002; 2002JP-00273125.

(GCDE) GC CORP.

Senpuku H, Masuzawa Y, Okada J;

WPI; 2004-284981/27.

Examining carries risk of person, involves using synthetic peptide having specific amino acid sequence as antigen, and measuring antibody value of secretory immunoglobulin A in human saliva against antigen.

Claim 1; Page 8; 11pp; English.

The present sequence represents a peptide which is derived from amino acid 361-386 of protein antigen serotype C (PAC). PAC is a surface protein antigen from Streptococcus mutans. The peptide is used as an antigen in the method of the invention. The specification describes a method for examining a carries risk of a person. The method involves using the present peptide as an antigen, and measuring an antibody value of a secretory immunoglobulin A in a human saliva against the antigen. The method accurately examines the carries risk from the difference of antibody value of human immunoglobulin to the antigen in a short time.

Sequence 26 AA;

Query Match 100.0%; Score 124; DB 8; Length 26;
 Best Local Similarity 100.0%; Pred. No. 1.1e-11;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAKATYEALKOYEADLAAYKANA 26
 |||||
 1 NAKATYEALKOYEADLAAYKANA 26

RESULT 2

ADO34199 standard; peptide; 26 AA.

AC ADO34199;

DT 26-AUG-2004 (first entry)

DE Protein Antigen serotype C, PAC, peptide fragment residues 361-386.

KM Dental caries; DRB1*; human leukocyte antigen, HLA;

KW secretory immunoglobulin A; antigen; protein antigen serotype C; PAC.

OS Synthetic.

PN EP1426765-A2.

PD 09-JUN-2004.

PF 18-NOV-2003; 2003EP-00026551.

PR 04-DEC-2002; 2002JP-00352466.

PA (GCDE) GC CORP.

PI Sempuku H, Masuzawa Y;

XX WPI; 2004-422608/40.

PT Examining caries risk, by identifying genotype of beta chain of DR gene

PS DB1sterisk in class II type of human leukocyte antigen gene group.

PS Claim 2; Page 7; 8pp; English.

XX The present invention relates to a method (M1) for examining dental

CC carries risk by identifying the genotype of the beta chain of the DR gene

CC (DRB1*) in a class II type of human leukocyte antigen (HLA) gene group.

CC In (M1), the identified genotype is compared with the caries risk which

CC had been previously identified, derived from an antibody value of a

CC secretory immunoglobulin A in human saliva against an antigen, in which

CC peptide ADO34199 is used as the antigen.

CC Sequence 26 AA;

QY 1 NAKATYEALKOYEADLAAYKANA 26
 |||||
 1 NAKATYEALKOYEADLAAYKANA 26

Db 1 NAKATYEALKOYEADLAAYKANA 26

QY 1 NAKATYEALKOYEADLAAYKANA 26
 |||||
 1 NAKATYEALKOYEADLAAYKANA 26

Db 1 NAKATYEALKOYEADLAAYKANA 26

QY 1 NAKATYEALKOYEADLAAYKANA 26
 |||||
 1 NAKATYEALKOYEADLAAYKANA 26

Db 1 NAKATYEALKOYEADLAAYKANA 26

QY 1 NAKATYEALKOYEADLAAYKANA 26
 |||||
 1 NAKATYEALKOYEADLAAYKANA 26

Db 1 NAKATYEALKOYEADLAAYKANA 26

QY 1 NAKATYEALKOYEADLAAYKANA 26
 |||||
 1 NAKATYEALKOYEADLAAYKANA 26

Db 1 NAKATYEALKOYEADLAAYKANA 26

XX Streptococcus mutans.
 OS
 XX Key Location/Qualifiers
 FH MISC-difference 618..650
 FT /note="amino acids 618-650 differ from the residues
 FT deduced from the nucleic acid sequence."

PN WO9623886-A1.

PD 08-AUG-1996.

PF 31-JAN-1996; 96WO-GB000207.

PR 31-JAN-1995; 95GB-00001826.

PA (UNME-) UNITED MEDICAL & DENTAL SCHOOLS GUYS.

PI Lehner T, Kelly C;

DR WPI, 1996-371434/37.

DR N-PSDB; AAT36122.

PT Nucleic acid encoding polypeptide for prevention or treatment of dental

PT caries - which stimulates T or B cell response, and/or adheres to tooth

PT in competition with Streptococcus mutans antigen I/II.

PS Disclosure; Page 45-46; 63pp; English.

XX Streptococcus mutans antigen I/II (AAW02098) is a 185 kDa cell surface

CC protein at least partly responsible for S. mutans adhesion to teeth. The

CC I/II antigen includes a series of overlapping T-cell, B-cell and adhesion

CC epitopes. Fragments (see also AAT36111-21) of the I/II antigen gene (see

CC also AAT36122) can be used to produce recombinant polypeptides (AAW02087-

CC 97) carrying such epitopes for use in vaccines for immunisation against

CC dental caries

CC Sequence 1561 AA;

QY 1 NAKATYEALKOYEADLAAYKANA 26
 |||||
 361 NAKATYEALKOYEADLAAYKANA 386

Db 361 NAKATYEALKOYEADLAAYKANA 386

QY 1 NAKATYEALKOYEADLAAYKANA 26
 |||||
 361 NAKATYEALKOYEADLAAYKANA 386

Db 361 NAKATYEALKOYEADLAAYKANA 386

QY 1 NAKATYEALKOYEADLAAYKANA 26
 |||||
 361 NAKATYEALKOYEADLAAYKANA 386

Db 361 NAKATYEALKOYEADLAAYKANA 386

QY 1 NAKATYEALKOYEADLAAYKANA 26
 |||||
 361 NAKATYEALKOYEADLAAYKANA 386

Db 361 NAKATYEALKOYEADLAAYKANA 386

QY 1 NAKATYEALKOYEADLAAYKANA 26
 |||||
 361 NAKATYEALKOYEADLAAYKANA 386

Db 361 NAKATYEALKOYEADLAAYKANA 386

QY 1 NAKATYEALKOYEADLAAYKANA 26
 |||||
 361 NAKATYEALKOYEADLAAYKANA 386

Db 361 NAKATYEALKOYEADLAAYKANA 386

QY 1 NAKATYEALKOYEADLAAYKANA 26
 |||||
 361 NAKATYEALKOYEADLAAYKANA 386

Db 361 NAKATYEALKOYEADLAAYKANA 386

QY 1 NAKATYEALKOYEADLAAYKANA 26
 |||||
 361 NAKATYEALKOYEADLAAYKANA 386

```

XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI: 2003-029926/02.
XX DR N-PSDB; AC488225.
XX
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX
XX PS Claim 25; SEQ ID NO 72279; 1766bp; English.
XX
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than S. aureus, S. typhimurium,
XX CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
XX CC the target prokaryotic essential genes. Note: The sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 1561 AA;
XX
XX Query Match 100.0%; Score 124; DB 6; Length 1561;
XX Best Local Similarity 100.0%; Pred. No. 1.5e-09;
XX Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 NAKATYEAALKQYEDLAAYKKANNA 26
XX Db ||||| ||||| ||||| ||||| |||||
XX 361 NAKATYEAALKQYEDLAAYKKANNA 386
XX
XX RESULT 5
XX ID AAR79643
XX AAR79643 standard; protein; 1566 AA.
XX
XX AC AAR79643;
XX
XX DT 27-MAR-1996 (first entry)
XX
XX DE Immunodominant antigen of Streptococcus sobrinus.
XX
XX KW Streptococci; enterococci; sera; antibodies; diagnosis; prophylaxis;
XX KW treatment; therapy; infection; endocarditis; septicemia.
XX
XX OS Streptococcus sobrinus.
XX

```

```

EN MN W09520658-A2.
XX
XX PD 03-AUG-1995.
XX
XX PF 30-JAN-1995; 95WO-GB000186.
XX
XX PR 28-JAN-1994; 94GB-00001689.
XX
XX PA (UTMA-) UNIV VICTORIA MANCHESTER.
XX
XX PI Burnie JP, Matthews RC;
XX
XX DR WPI: 1995-275443/36.
XX DR N-PSDB; AA097812.
XX
XX PT New purified bacterial protein - used to develop prods. for the diagnosis
XX PT and treatment of infections due to streptococci and enterococci.
XX
XX PS Example 5; Page 48-54; 92pp; English.
XX
XX CC Purified bacterial proteins expressed during infection by streptococci or
XX CC enterococci and isolated from human sera may be used to develop products
XX CC e.g. antibodies, for use in the diagnosis, prophylaxis and treatment of
XX CC infections caused by these organisms, especially endocarditis and
XX CC septicemia. This sequence is the immunodominant antigen of Streptococcus
XX CC sobrinus.
XX
XX SQ Sequence 1566 AA;
XX
XX Query Match 62.1%; Score 77; DB 2; Length 1566;
XX Best Local Similarity 64.0%; Pred. No. 0.021;
XX Matches 16; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
XX
XX QY 2 AKATYEAALKQYEDLAAYKKANNA 26
XX Db ||||| ||||| ||||| ||||| |||||
XX 365 AKATYEAALKQYEDLAAYKKANNA 389
XX
XX RESULT 6
XX ID AAP80522
XX AAP80522 standard; peptide; 19 AA.
XX
XX AC AAP80522;
XX
XX DT 25-MAR-2003 (revised)
XX DT 31-OCT-2002 (revised)
XX DT 19-NOV-1990 (first entry)
XX
XX DE Synthetic peptide identical to fragments of Streptococcus mutans antigen
XX DE I, antigen I/II or antigen X.
XX
XX KW Anticaries; vaccine.
XX
XX OS Synthetic.
XX
XX PN EP280576-A.
XX
XX PD 31-AUG-1988.
XX
XX PE 26-FEB-1988; 88EP-00301704.
XX
XX PR 27-FEB-1987; 87GB-00004647.
XX
XX PA (GUY-) GUY'S & ST THOMAS.
XX PA (JOHJ) JOHNSON & JOHNSON CONSUM.
XX PA (UNME-) UNITED MED & DENTAL.
XX
XX PI Lehner T, Haron JA, Freidmann N;
XX
XX DR WPI: 1988-244018/35.
XX
XX PT Synthetic anticaries peptide - capable of provoking formation of
XX PT antibodies in vivo that recognise Streptococcus mutans.
XX

```

xx	
ps	Claim 7; Page 13; 21pp; English.
cc	
cc	The novel synthetic peptide has an AA sequence identical to fragments of
cc	antigen I or antigen I/I1 or antigen X of S. mutans. The peptide directs
cc	production of specific anti- S. mutans antibodies and stimulates helper T-
cc	lymphocyte proliferation to combat dental caries. The peptide may be used
cc	for diagnostic purposes or for passive immunization against S. mutans
cc	infection. (See also AAP80516, AAP80518, AAP80520, AAP80517, AAP80521,
cc	AAP80523, AAP80524, AAP80525, AAP80526). (Updated on 31-OCT-2002 to add
cc	missing OS field.) (Updated on 25-MAR-2003 to correct PR field.) (Updated
cc	on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct
cc	PI field.)
cc	
sq	Sequence 19 AA;
qy	
Db	
2 AKATYEALIKQYEAADLAA 19	
1 AKADYEAKLAQYEXDLAA 18	
RESULT 7	
ID ABB70681	51.6%; Score 64; DB 1; Length 19;
ID ABB70681 standard; protein; 1296 AA.	Best Local Similarity 77.8%; Pred.No.0.01;
XX ABB70681;	Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
AC ABB70681;	
LT 26-MAR-2002 (first entry)	
DE Drosophila melanogaster polypeptide SEQ ID NO 38835.	
KW Drosophila; developmental biology; cell signalling; insecticide;	
KM pharmaceutical.	
OS Drosophila melanogaster.	
PN MO200171042-AA2.	
PD 27-SEP-2001.	
PF 23-MAR-2001; 2001WO-US009231.	
PR 23-MAR-2000; 2000US-0191637P.	
PR 11-JUL-2000; 2000US-00614150.	
XX (PEKE) PE CORP NY.	
PA Venter JC, Adams M, Li PWD, Myers EW;	
P1 WPI; 2001-656860/75.	
DR N-PESDB; ABLL4784.	
PT New isolated nucleic acid detection reagent for detecting 1000 or more	
PT genes from Drosophila and for elucidating cell signaling and cell-cell	
PT interactions.	
XX	
PS Disclosure; SEQ ID NO 38835; 21pp + Sequence Listing; English.	
CC	
CC The invention relates to an isolated nucleic acid detection reagent	
CC capable of detecting 1000 or more genes from Drosophila. The invention is	
CC useful in developmental biology and in elucidating cell signaling and	
CC cell-cell interactions in higher eukaryotes for the development of	
CC insecticides, therapeutics and pharmaceutical drugs. The invention	
CC discloses genomic DNA sequences (ABLL16176-ABLL30511), expressed DNA	
CC sequences (ABL01840-ABLL16175) and the encoded proteins (ABB57737-	
CC ABB120722). The sequence data for this patent did not form part of the	
CC printed specification, but was obtained in electronic format directly	
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences	
XX	

Query Match	Best Local Similarity	Score	DB	Length
12; Conservative	57.1%	57	4	1296
Matches		4	Mismatches	5
			Indels	0
			Gaps	0

Seq	Sequence	1296 AA
Qy	6 YEALAKOYEADLAAYKKNAA 26	
Db	481 FEALFKOYEAVTASLKQLNFA 501	

Result	ID	ADO000993	standard; protein; 1296 AA.
AC	ADO000993;		
XX	01-JUL-2004	(first entry)	
XX	Fruit fly AD-related protein CG16972.		
XX	Fruit fly; Alzheimer's disease; Gamma secretase; Psn gene; P-element; EP;		
XX	APPL-SV; Amyloid precursor-like protein; APP;		
XX	suppressor of hairless transcription factor; Su(H);		
XX	VP16 activation domain; dementia; memory loss; language deterioration;		
XX	impaired visuospatial skill.		
XX	Drosophila melanogaster.		
XX	US2004067535-A1.		
XX	08-APR-2004.		
XX	03-OCT-2002; 2002US-00263929.		
XX	03-OCT-2002; 2002US-00263929.		
XX	(LIFE-) LIFE SCI DEV CORP.		
XX	Kim J, Galant R.		
XX	WPI; 2004-355296/33.		
XX	N-PSDB; ADO000891.		
XX	Identifying compound by exposing cell that expresses gene having		
XX	enhancing or suppression effect on APPL-SV phenotype to agent,		
XX	identifying modulation of Alzheimer's disease (AD), regulation of gene or		
XX	protein expression with AD.		
XX	Claim 18; SEQ ID NO 131; 185pp; English.		

Seq	Sequence	1296 AA
XX	The invention relates to identifying a compound comprising exposing cell	
XX	expressing gene 1 having enhancing or suppression effect on an APPL-SV	
XX	phenotype (a transgenic fruit fly expressing the Amyloid precursor-like	
XX	protein, APP, as a fusion protein with the suppressor of hairless	
XX	transcription factor, Su(H) and VP16 activation domain. The fusion	
XX	protein is cleaved by gamma secretase (encoded by the Psn gene) to	
XX	release the Su(H)-VP16 which affects wing vein development. Genes	
XX	affecting Psn expression/activity were screened by crossing the APPL-SV	
XX	line with an EP P-element insertion library, and the DNA recovered from	
XX	the appropriate EP strain and sequenced) chosen from ADO000863-ADO00864,	
XX	being the identified fruit fly genes affecting APP processing and their	
XX	mammalian homologues, identifying modulation of Alzheimer's disease (AD)	
XX	symptom, regulation of biological pathway, gene expression or protein	
XX	function associated with AD relative to cell in absence of agent. Also	
XX	included are regulating AD (involves providing a subject with AD or	
XX	symptoms of AD and an agent that changes the expression of a gene	
XX	described above or changes the activity of a polypeptide having a sequence	
XX	chosen from ADO000965-ADO01066, and treating the subject with the agent)	
XX	and a composition (comprising a nucleic acid encoding a polypeptide	
XX	a detailed above or an expression vector comprising the nucleic acid or a	
XX	host cell comprising the expression vector or an antisense	
XX	oligonucleotide that hybridizes under stringent conditions to the nucleic	

ID ADR73469 standard; peptide; 16 AA.
XX
XX ADR73469;
AC
XX
DT 09-SEP-2004 (first entry)
XX
DB Streptococcus sobrinus B cell epitope of gene AgI/II, SEQ ID 82.
XX
XX transgenic animal; Hepatitis B virus nucleocapsid core protein; Hbc;
KW enhanced stability; hepatotropic; virucide; immunology;
KW protein engineering; immunogen; vaccine; Hepatitis B infection.
XX
OS Streptococcus sobrinus.
XX
PN WO2004053091-A2.
XX
XX 24-JUN-2004.
PD
XX 10-DEC-2003; 2003WO-US039164.
PF
XX 10-DEC-2002; 2002US-0432123P.
PR
XX 10-DEC-2002; 2002US-0432123P.
XX
XX (APOV-) APOVIA INC.
PA
XX Lyons K, Birkett AJ, Haron JA;
PI
XX WPI; 2004-468859/44.
DR
XX
PT New recombinant chimer hepatitis B core (Hbc) protein molecules useful in
PT the fields of immunology and protein engineering, in particular as an
PT immunogen in a vaccine for Hepatitis B infections.
XX
XX
XX Disclosure; SEQ ID NO 82; 338pp; English.
PS
XX The invention relates to a novel recombinant chimeric Hepatitis B virus
CC nucleocapsid (core) protein (Hbc), up to 600 or 380 amino acid residues
CC in length. The chimeric protein is engineered for both enhanced stability
CC of self-assembled particles and the substantial absence of nucleic acid
CC binding by the particles. The invention further comprises: a recombinant
CC Hbc protein chimeric molecule that has a length of 135-365 amino acid
CC residues and contains four peptide-linked amino acid residue sequence
CC domains from the N-terminus that are denominated Domains I, II, III and
CC IV. The invention also provides nucleic acids, polypeptides, host cells,
CC vectors and transgenic animals used in the methods of the invention. The
CC chimeric compositions of the invention have hepatotropic and virucide
CC activities. The methods and compositions of the present invention are
CC useful in the fields of immunology and protein engineering, in particular
CC for using a chimeric hepatitis B virus nucleocapsid protein as an
CC immunogen in a vaccine for Hepatitis B infections. This sequence
CC represents a Hepatitis B virus nucleocapsid (core) protein related
CC polypeptide of the invention.
XX
XX Sequence 16 AA;
SQ
Query Match 45.2%; Score 56; DB 8; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.14;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 2 AKATYEALKOYEADL 17
DB 1 AKADYEAKLAQYERDL 16
RESULT 14
ID ADR12762 standard; peptide; 16 AA.
XX
XX ADR12762;
AC
XX 04-NOV-2004 (first entry)
DT
XX S. sobrinus AgI/II B cell epitope #2.
DE
XX

KW HBV, chronic hepatitis; Hbc; nucleocapsid core protein; vaccine;
KW immunogen; immunogenic epitope; T cell; B cell; CD8+ cell; CD4+ cell;
KW cytotoxic T lymphocyte; toll-like receptor 4; toll-like receptor 9;
KW TLR-4; TLR-9; epitope.
XX
XX Streptococcus sobrinus.
OS
XX US2004156863-A1.
PN
XX 12-AUG-2004.
PD
XX 01-OCT-2003; 2003US-00677074.
PF
XX 21-FEB-2002; 2002US-00080299.
PR 21-FEB-2002; 2002US-00082014.
PR 21-FEB-2003; 2003US-00372076.
XX
XX (PAGE/) PAGE M.
PA (FRIE/) FRIEDE M.
PA (SCHM/) SCHMIDT A E.
PA (STOB/) STOBBER D.
XX
XX Page M, Friede M, Schmidt AE, Stober D;
PI
XX WPI; 2004-603322/58.
DR
XX
XX Treating chronic hepatitis, by administering vaccine comprising
PT immunogenic particles having recombinant hepatitis B core chimeric
PT protein molecules, that stimulates T cell, to patient chronically
PT infected with hepatitis B virus.
XX
XX
XX Disclosure; SEQ ID NO 69; 117pp; English.
PS
XX The invention relates to treating chronic hepatitis, by administering a
CC vaccine comprising immunogenic particles having recombinant hepatitis B
CC core (Hbc) chimeric protein molecules (where truncated Hbc molecules are
CC linked N-terminally or C-terminally to an immunogenic epitope), that
CC stimulate T cell production, to a patient chronically infected with
CC hepatitis B virus, and maintaining patient for time sufficient to induce
CC T cells activated against Hbc. The chimeric proteins are still capable
CC self-assembling into particles upon expression in a host cell and are
CC substantially free of binding to nucleic acids, and the particles display
CC enhanced stability. Also included is enhancing (M2) the production of one
CC or more of gamma-producing CD8+, CD4+ T cells and cytotoxic T lymphocytes
CC against hepatitis B virus, involving administering to a patient
CC chronically infected with hepatitis B virus, a T cell-stimulating amount
CC of a vaccine comprising immunogenic particles dissolved or dispersed in a
CC diluent containing one or both of an agonist of toll-like receptor 4 and
CC receptor 9 (TLR-4 and TLR-9), the immunogenic particles comprising Hbc
CC chimeric protein molecules and maintaining the patient for a sufficient
CC time to induce T cells activated against Hbc. The immunogenic epitopes
CC may be B cell or T cell epitopes. The chimeric vaccine is useful for
CC treating a patient chronically infected with hepatitis B virus. The
CC present sequence is a B cell epitope suitable for inclusion in the
CC chimeric protein of the invention.
XX
XX Sequence 16 AA;
SQ
Query Match 45.2%; Score 56; DB 8; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.14;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 2 AKATYEALKOYEADL 17
DB 1 AKADYEAKLAQYERDL 16
RESULT 15
ID AAY29700 standard; protein; 17 AA.
XX
XX AAY29700;
AC
XX AAY29700;
XX


```

DT 08-NOV-1999 (first entry)
XX
DE Streptococcus sobrinus antigen AgI/II polypeptide hapten.
XX
KW Human hepatitis B core protein; Hbc; modified; immunodominant;
XX nucleocapsid protein; vaccine; T cell epitope.
OS Streptococcus sobrinus.
XX
PN WO9940934-A1.
XX
PD 19-AUG-1999.
XX
PF 11-FEB-1999; 99WO-US003055.
XX
PR 12-FEB-1998; 98US-0074537P.
XX
PA (IMMU-) IMMUNE COMPLEX CORP.
XX
PI Birkett AJ;
XX
DR WPI, 1999-527340/44.
XX
PT Conjugate of hepatitis B core protein, modified to increase reactivity
PT with hapten, used to raise antibodies against the hapten, e.g. in
PT vaccines.
XX
PS Example 3, Page 38, 128pp; English.
XX
XX
CC The present invention describes a conjugate (A) comprising a
CC strategically modified hepatitis B core (Hbc) protein (I) attached to a
CC hapten, where (I) includes amino acids (aa) 10-140 of the wild type Hbc
CC 183 aa sequence (given in AA29674) and additionally has an insert (II)
CC in the region corresponding to aa's 50-100, where the insert is of 1 to
CC about 40 aa's and contains a chemically reactive aa residue linked to the
CC hapten. A vaccine containing (A), optionally in the form of particles, is
CC used to induce a protective antibody response against the pathogen from
CC which the hapten is derived, in humans or other animals. These pathogens
CC may be bacteria, viruses, rickettsia or protozoa. Insertion of (II)
CC overcomes the low reactivity of aa side chains in native Hbc protein,
CC increasing the reactivity with hapten and resulting in conjugates of
CC improved immunogenicity. Modified Hbc can be derivatised in the form of
CC particles by well-defined chemical methods, and is unlikely to cause
CC immunological side-effects. AA29675 to AA29735 represent polypeptide
CC haptens used in an example from the present invention
XX
SQ Sequence 17 AA;

```

```

Query Match 45.2%; Score 56; DB 2; Length 17;
Best Local Similarity 75.0%; Pred No. 0.15;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 2 AKATYEALKQYEADL 17
   ||| ||| ||| |||
Db 1 AKADYEAKLAQYEKDL 16

```

```

Search completed: October 14, 2005, 14:48:09
Job time : 73 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2005, 14:32:07 ; Search time 16 Seconds

(without alignments)
156.352 Million cell updates/sec

Title: US-10-714-852-1

Sequence: 1 NAKATYEALKKQYEDLAIVKXANAA 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR.79:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	100.0	1556	2 A60988	saliva-interacting
2	124	100.0	1561	1 S06839	surface antigen sp
3	124	100.0	1565	2 S04729	surface antigen pa
4	107	86.3	1473	2 A35186	salivary agglutini
5	93	75.0	365	2 S70994	cell surface adhes
6	79	63.7	1528	2 A60338	surface antigen A
7	77	62.1	1566	2 A43607	cell surface anti-g
8	55	44.4	1136	2 AB1581	different protein
9	54	43.5	982	2 T18576	serine-threonine k
10	54	43.5	1256	1 A43829	muramidase-release
11	52	41.9	1136	2 AH1227	different proteins
12	51	41.1	545	2 T21253	hypothetical prote
13	51	41.1	689	2 T34025	hypothetical prote
14	50	39.9	1620	2 S61535	nucleotide-binding
15	49.5	39.9	370	2 A45981	peptidylprolyl iso
16	49.5	39.9	370	2 A46579	estrogen receptor-
17	49	39.5	239	2 AE1317	hypothetical prote
18	49	39.5	757	2 UC7726	(1->4)-alpha-D-glu
19	49	39.5	863	2 S37040	parmyosin - tapew
20	48.5	39.1	234	2 F70416	hypothetical prote
21	48.5	39.1	372	2 E81290	probable capsule p
22	48	38.7	244	2 I79263	trial protein - Esc
23	48	38.7	274	1 BVBRCE	cyas protein - Bor
24	47.5	37.9	274	2 D95339	hypothetical prote
25	47	37.9	118	2 F97144	hypothetical prote
26	47	37.9	309	2 A43583	adhesin B precursor
27	47	37.9	360	2 F96618	RNA binding protei
28	47	37.9	364	2 D95364	hypothetical prote
29	47	37.9	381	2 T52461	RNA binding protei

30	47	37.9	891	2 G41662	130K surface exclu
31	47	37.9	1634	2 T26517	hypothetical prote
32	47	37.9	2032	2 I39917	hypothetical prote
33	46.5	37.5	429	2 F87282	conserved hypochet
34	46	37.1	181	2 AE3291	5-formyltetrahydro
35	46	37.1	310	2 T11551	adhesin - Streptoc
36	46	37.1	319	2 AG0742	high-affinity zinc
37	46	37.1	466	2 E86903	hypothetical prote
38	46	37.1	501	2 T39801	hypothetical sn3-c
39	46	37.1	856	2 T34491	hypothetical prote
40	45	36.3	132	2 B75339	hypothetical prote
41	45	36.3	185	2 B82776	hypothetical prote
42	45	36.3	199	2 E46449	hypothetical prote
43	45	36.3	227	2 S04133	photosystem I chai
44	45	36.3	327	2 D82678	transcription regu
45	45	36.3	539	2 S55190	hypothetical prote

ALIGNMENTS

RESULT 1
saliva-interacting protein precursor - Streptococcus mutans (serotype f, strain OM2175
N, Alternate names: cell surface antigen spap homolog, SR protein
C, Species: Streptococcus mutans
C, Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C, Accession: A60988; A60660
R, Ogier, J.A.; Scholler, M.; Lepoint, Y.; Pini, A.; Sommer, P.; Klein, J.P.
FEMS Microbiol. Lett. 68, 223-228, 1990
A, Title: Complete nucleotide sequence of the sr gene from Streptococcus mutans OM2 175
A, Reference number: A60988
A, Accession: A60988
A, Molecule type: DNA
A, Residues: 1-1556 <OG1>
A, Cross-references: UNIPROT:Q07270; GB:X53647; NID:Q297828; PIDN:CAA37697.1; PID:Q2978
R, Ogier, J.A.; Wachmann, D.; Scholler, M.; Lepoint, Y.; Klein, J.P.
Arch. Oral Biol. 35(Suppl.), 285-315, 1990
A, Title: Molecular characterization of the gene sr of the saliva interacting protein f
A, Reference number: A60660; WUID:91207142; PMID:2088234
A, Accession: A60660
A, Status: nucleic acid sequence not shown; not compared with conceptual translation
A, Molecule type: DNA
A, Residues: 1,1171,'S',1173-1556 <OG2>
C, Comment: This protein probably plays a role in adherence to the tooth surface.
C, Genetics:
A, Gene: sr
C, Superfamily: surface antigen spap
C, Keywords: duplication; glycoprotein; surface antigen
F, 1-38/Domain: signal sequence #status predicted <StC>
F, 39-1556/Product: saliva-interacting protein #status predicted <MAT1>
F, 170-194/Region: alanine-rich repeat
F, 195-219/Region: alanine-rich repeat
F, 220-244/Region: alanine-rich repeat
F, 245-269/Region: alanine-rich repeat
F, 277-301/Region: alanine-rich repeat
F, 302-326/Region: alanine-rich repeat
F, 327-351/Region: alanine-rich repeat
F, 359-383/Region: alanine-rich repeat
F, 384-408/Region: alanine-rich repeat
F, 409-427/Region: alanine-rich repeat
F, 441-465/Region: alanine-rich repeat
F, 482-890/Region: proline-rich repeat
F, 891-929/Region: proline-rich repeat
F, 930-968/Region: proline-rich repeat
Query Match 100.0%; Score 124; DB 2; Length 1556;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAKATYEALKKQYEDLAIVKXANAA 26
DB 362 NAKATYEALKKQYEDLAIVKXANAA 387

RESULT 2

S06839

surface antigen spap precursor - Streptococcus mutans

N/Alternate names: antigen I/II

C/Species: Streptococcus mutans

C/Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 09-Jul-2004

C/Accession: S06839; A60339; A60661

R/Kelly, C.; Evans, P.; Bergmeier, L.; Lee, S.F.; Progluske-Fox, A.; Harris, A.C.; Altke
FEBS Lett. 258, 127-132, 1989

A/Title: Sequence analysis of the cloned streptococcal cell surface antigen I/II.

A/Reference number: S06839; MUID:90076473; PMID:2687020

A/Accession: S06839

A/Molecule type: DNA

A/Residues: 1-1561 <REL>

A/Cross-references: UNIPROT:P2504; EMBL:X17390; NID:947266; PIDN:CA35253.1; PID:947267

A/Experimental source: serotype C

A/Note: parts of this sequence, including the amino ends of the mature proteins, were CD

R/Kelly, C.; Evans, P.; Ma, J.K.C.; Bergmeier, L.A.; Taylor, W.; Brady, L.J.; Lee, S.F.;

Arch. Oral Biol. 35(Suppl.), 33S-38S, 1990

A/Title: Sequencing and characterization of the 185 kDa cell surface antigen of Streptoc

A/Reference number: A60661; MUID:91207143; PMID:1982405

A/Contents: annotation

R/Ma, J.K.C.; Kelly, C.G.; Munro, G.; Whitley, R.A.; Lehner, T.

Intect. Immun. 59, 2686-2694, 1991

A/Title: Conservation of the gene encoding streptococcal antigen I/II in oral streptococ

A/Reference number: A60339; MUID:91310321; PMID:1855988

A/Accession: A60339

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1084-1189 <MAA>

C/Genetics:

A/Function:

A/Description: probably plays a role in adherence to the tooth surface.

C/Superfamily: surface antigen spap

C/Keywords: duplication; glycoprotein; transmembrane protein

F/1-38/Domain: signal sequence #status predicted <SIG>

F/39-1561/Product: surface antigen spap.I #status predicted <MAT1>

F/159-193/Domain: spap alanine-rich repeat <AR1>

F/194-218/Domain: spap alanine-rich repeat <AR2>

F/219-243/Domain: spap alanine-rich repeat <AR3>

F/244-268/Domain: spap alanine-rich repeat <AR4>

F/276-300/Domain: spap alanine-rich repeat <AR5>

F/301-325/Domain: spap alanine-rich repeat <AR6>

F/326-350/Domain: spap alanine-rich repeat <AR7>

F/358-382/Domain: spap alanine-rich repeat <AR8>

F/383-407/Domain: spap alanine-rich repeat <AR9>

F/408-426/Domain: spap alanine-rich repeat <AR10>

F/440-464/Domain: spap alanine-rich repeat <AR11>

F/847-885/Domain: spap proline-rich repeat <PR1>

F/886-924/Domain: spap proline-rich repeat <PR2>

F/925-963/Domain: spap proline-rich repeat <PR3>

F/997-1561/Product: surface antigen spap.II #status predicted <MAT2>

Query Match

Best Local Similarity 100.0%; Score 124; DB 1; Length 1561;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 NAKATYEAALKOYEADLAIVKKNAA 26

Db 361 NAKATYEAALKOYEADLAIVKKNAA 386

RESULT 3

S04729

surface antigen pac precursor - Streptococcus mutans

C/Species: Streptococcus mutans

C/Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004

C/Accession: S04729

R/Okahashi, N.; Sasakawa, C.; Yoshikawa, M.; Hamada, S.; Koga, T.

Mol. Microbiol. 3, 673-678, 1989

A/Title: Molecular characterization of a surface protein antigen gene from serotype c ;

A/Reference number: S04729; MUID:89343654; PMID:2761390

A/Accession: S04729

A/Molecule type: DNA

A/Residues: 1-1565 <OKA>

A/Cross-references: UNIPROT:P11557; EMBL:X14490; NID:947247; PIDN:CA32652.1; PID:9472-

A/Note: part of this sequence, including the amino end of the mature protein, was conf.

C/Genetics:

A/Function: pac

C/Superfamily: surface antigen spap

C/Keywords: surface antigen; transmembrane protein

F/38/Domain: signal sequence #status predicted <SIG>

F/39-1565/Product: surface antigen pac #status experimental <MAT>

Query Match

Best Local Similarity 100.0%; Score 124; DB 2; Length 1565;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 NAKATYEAALKOYEADLAIVKKNAA 26

Db 361 NAKATYEAALKOYEADLAIVKKNAA 386

RESULT 4

A35186

salivary agglutinin receptor precursor - Streptococcus sanguis

C/Species: Streptococcus sanguis

C/Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 09-Jul-2004

C/Accession: A35186

R/Demuth, D.R.; Golub, E.E.; Malamud, D.

J. Biol. Chem. 265, 7120-7126, 1990

A/Title: Streptococcal-host interactions. Structural and functional analysis of a Stre

A/Reference number: A35186; MUID:90236997; PMID:2185241

A/Accession: A35186

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1473 <DEM>

A/Cross-references: UNIPROT:P16952; GB:J05418

C/Superfamily: surface antigen spap

Query Match

Best Local Similarity 86.3%; Score 107; DB 2; Length 1473;

Matches 21; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 2 AKATYEAALKOYEADLAIVKKN 24

Db 280 AKATYEAALKOYEADLAIVKKN 302

RESULT 5

S70994

cell surface adhesin B - Streptococcus gordonii (fragment)

C/Species: Streptococcus gordonii

C/Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 09-Jul-2004

C/Accession: S70994

R/Demuth, D.R.; Duan, Y.; Brooks, W.; Holmes, A.R.; McNab, R.; Jenkinson, H.F.

Mol. Microbiol. 20, 403-413, 1996

A/Title: Tandem genes encode cell-surface polypeptides SspA and SspB which mediate adh

A/Reference number: S70994; MUID:96310377; PMID:8733238

A/Accession: S70994

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-365 <DEM>

A/Cross-references: UNIPROT:Q54186; EMBL:U40027; NID:91100974; PIDN:AC44102.1; PID:91-

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995

C/Genetics:

A/Function: sspB

C/Superfamily: surface antigen spap

Query Match

Best Local Similarity 75.0%; Score 93; DB 2; Length 365;

Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 AKATYEAALKQYEDLAAVKKA 23
 Db 279 AKAYDAAKQYEDLAAIKKA 300

RESULT 6

A60338
 surface antigen A - Streptococcus sobrinus (strain 6715, serotype g)
 N:Alternate names: streptococcal antigen I/II
 C:Species: Streptococcus sobrinus

C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 20-Aug-1999

C/Accession: A60338; S21811

R:LaPolia, R.O.; Haron, U.A.; Kelly, C.G.; Taylor, W.R.; Bohart, C.; Hendricks, M.; Pyat

Infect. Immun. 59, 2677-2685, 1991

A>Title: Sequence and structural analysis of surface protein antigen I/II (Spa) of Stre

A:Reference number: A60338; MUID:91310320; PMID:1855987

A:Accession: A60338

A:Molecule type: DNA

A:Residues: 1-1528 <LAP>

A:Cross-references: GB:X57841; NID:G47619; PIDN:CAA40973.1; PID:G47620

C:Comment: This protein appears to be an adhesin involved in attachment of the bacterium

C:Genetics:

A:Gene: spa

C:Superfamily: surface antigen spa

C:Keywords: surface antigen; tandem repeat

Query Match 63.7%; Score 79; DB 2; Length 1528;

Best Local Similarity 65.4%; Pred. No. 0.0049; Indels 0; Gaps 0;

Matches 17; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 NAKATYEAALKQYEDLAAVKKANA 26

Db 292 SAKADYEAALKQYEDLAAQAGNNA 317

RESULT 7
 cell surface antigen Spa precursor - Streptococcus sobrinus (strain MT3791)
 C/Species: Streptococcus sobrinus

C>Date: 12-Jan-1993 #sequence_revision 12-Jan-1993 #text_change 09-Jul-2004

C/Accession: A43607

R:Tokuda, M.; Okanashi, N.; Takahashi, I.; Nakai, M.; Nagaoka, S.; Kawagoe, M.; Koga, T.

Infect. Immun. 59, 3309-3312, 1991

A>Title: Complete nucleotide sequence of the gene for a surface protein antigen of Stre

A:Reference number: A43607; MUID:91348886; PMID:1840575

A:Accession: A43607

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1566 <TOX>

A:Cross-references: UNIPROT:Q55308; GB:D90354; NID:G217035; PIDN:BAA14368.1; PID:G217036

C:Superfamily: surface antigen spa

C:Keywords: surface antigen

Ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlan

A>Title: Comparative genomics of *Listeria species*.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AB1581

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1136 <GLA>

A:Cross-references: UNIPROT:Q92C14; GB:AL592022; PIDN:CAC96418.1; PID:G16413646; GSPDB

A:Experimental source: strain Clj11262

C:Genetics:

A:Gene: lln1187

Query Match 44.4%; Score 55; DB 2; Length 1136;

Best Local Similarity 50.0%; Pred. No. 9.6;

Matches 13; Conservative 7; Mismatches 4; Indels 2; Gaps 1;

OY 2 AKATYEAALKQYEDLAA--VKKANA 25

Db 303 AKASYDAKIKQSEETRAGEALANA 328

RESULT 9
 serine-threonine kinase - *Caenorhabditis elegans*
 C/Species: *Caenorhabditis elegans*

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C/Accession: T18576

R:COPE, M.J.T.V.; Kendrick-Jones, A.

submitted to the EMBL Data Library, July 1995

A:Reference number: Z18984

A:Accession: T18576

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-982 <BAR>

A:Cross-references: UNIPROT:P46549; EMBL:U32275; PIDN:AAA75370.1; GSPDB:GN00021; CESP:

A:Experimental source: strain N2 Bristol

C:Genetics:

A:Gene: CESP:slu

A:Map position: 3

Query Match 43.5%; Score 54; DB 2; Length 982;

Best Local Similarity 52.4%; Pred. No. 11;

Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 3 KATYEAALKQYEDLAAVKKA 23

Db 651 RSTYENAKKEVKADLRVKDA 671

RESULT 10
 muramidase-released protein precursor - *Streptococcus suis* (type 2, strain D282)

C/Species: *Streptococcus suis*

C>Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004

C/Accession: A43829; S21549

R:Smilh, H.E.; Vecht, U.; Gielkens, A.L.; Smits, M.A.

Infect. Immun. 60, 2361-2367, 1992

A>Title: Cloning and nucleotide sequence of the gene encoding the 136-kilodalton surfa

A:Reference number: A43829; MUID:92267650; PMID:1587602

A:Accession: A43829

A:Molecule type: DNA

A:Residues: 1-1256 <SMI>

A:Cross-references: UNIPROT:P32653; EMBL:X64450; NID:G47550; PIDN:CAA45781.1; PID:G475

C:Superfamily: muramidase-released protein

C:Keywords: tandem repeat; transmembrane protein

ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlan

A>Title: Comparative genomics of *Listeria species*.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AB1581

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1136 <GLA>

A:Cross-references: UNIPROT:Q92C14; GB:AL592022; PIDN:CAC96418.1; PID:G16413646; GSPDB

A:Experimental source: strain Clj11262

C:Genetics:

A:Gene: lln1187

Query Match 44.4%; Score 55; DB 2; Length 1136;

Best Local Similarity 50.0%; Pred. No. 9.6;

Matches 13; Conservative 7; Mismatches 4; Indels 2; Gaps 1;

OY 2 AKATYEAALKQYEDLAA--VKKANA 25

Db 303 AKASYDAKIKQSEETRAGEALANA 328

RESULT 9
 serine-threonine kinase - *Caenorhabditis elegans*
 C/Species: *Caenorhabditis elegans*

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C/Accession: T18576

R:COPE, M.J.T.V.; Kendrick-Jones, A.

submitted to the EMBL Data Library, July 1995

A:Reference number: Z18984

A:Accession: T18576

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-982 <BAR>

A:Cross-references: UNIPROT:P46549; EMBL:U32275; PIDN:AAA75370.1; GSPDB:GN00021; CESP:

A:Experimental source: strain N2 Bristol

C:Genetics:

A:Gene: CESP:slu

A:Map position: 3

Query Match 43.5%; Score 54; DB 2; Length 982;

Best Local Similarity 52.4%; Pred. No. 11;

Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 3 KATYEAALKQYEDLAAVKKA 23

Db 651 RSTYENAKKEVKADLRVKDA 671

RESULT 10
 muramidase-released protein precursor - *Streptococcus suis* (type 2, strain D282)

C/Species: *Streptococcus suis*

C>Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004

C/Accession: A43829; S21549

R:Smilh, H.E.; Vecht, U.; Gielkens, A.L.; Smits, M.A.

Infect. Immun. 60, 2361-2367, 1992

A>Title: Cloning and nucleotide sequence of the gene encoding the 136-kilodalton surfa

A:Reference number: A43829; MUID:92267650; PMID:1587602

A:Accession: A43829

A:Molecule type: DNA

A:Residues: 1-1256 <SMI>

A:Cross-references: UNIPROT:P32653; EMBL:X64450; NID:G47550; PIDN:CAA45781.1; PID:G475

C:Superfamily: muramidase-released protein

C:Keywords: tandem repeat; transmembrane protein

F:1232-1248/Domain: transmembrane #status predicted <TM1>

Query Match 43.5%; Score 54; DB 1; Length 1256;
Best Local Similarity 56.5%; Pred. No. 15;
Matches 13; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 4 ATYEALQYEAADLAAYKANA 26
Db 144 ADSEATVEQYEAQYAAVAVKANA 166

RESULT 11

AH1227 different proteins homolog lmo1224 [imported] - listeria monocytogenes (strain EGD-e)

C/Species: listeria monocytogenes
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C/Accession: AH1227

R/Gasser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgey, O.; Entian, K.D.; Fsihi, H.; D.; Jones, L.M.; Karet, U.

Science 294, 849-852, 2001

A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maicournam, A.; Maok, C.; Schluerer, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of listeria species

A/Reference number: AB1077; MUID:21537279; PMID:11679669

A/Accession: AH1227

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1136 <GLA>

A/Cross-references: UNIPROT:O8Y7P9; GB:NC_003210; PIDN:CAC9302.1; PID:G16410640; GSPDB: A/Experimental source: strain EGD-e

C/Genetics: A/Gene: lmo1224

Query Match 41.9%; Score 52; DB 2; Length 1136;
Best Local Similarity 39.4%; Pred. No. 26;
Matches 13; Conservative 8; Mismatches 4; Indels 8; Gaps 1;

Qy 2 AKATYEALQYEAADLAAYKANA 26
Db 303 ARASYDAKIKQSEAEINAGAEALNAKKQIDAA 335

RESULT 12

T21253 hypothetical protein F22D6.3 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C/Accession: T21253

R/Wilkinson, J.

submitted to the EMBL Data Library, April 1996

A/Reference number: Z19397

A/Accession: T21253

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-545 <WIL>

A/Cross-references: UNIPROT:Q19722; EMBL:Z71262; PIDN:CAA95808.1; GSPDB:GN00019; CESP:F2

A/Experimental source: clone F22D6

C/Genetics: A/Gene: CESP:F22D6.3

A/Map position: 1

A/Intons: 28/3; 318/2; 376/3; 430/3

C/Superfamily: lysine-tRNA ligase

Query Match 41.1%; Score 51; DB 2; Length 545;
Best Local Similarity 45.5%; Pred. No. 17;
Matches 10; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 2 AKATYEALQYEAADLAAYKKA 23
Db 55 AKAAIKKAVKKYEAEEVKKLEKA 76

RESULT 13

T34025

hypothetical protein C32F10.5 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 16-Aug-2004

C/Accession: T34025

R/Sammons, L.; Wohldmann, P.

submitted to the EMBL Data Library, April 1997

A/Description: The sequence of C. elegans cosmid C32F10.

A/Reference number: Z21465

A/Accession: T34025

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-689 <SAM>

A/Cross-references: UNIPROT:O01663; EMBL:AF000195; PIDN:AAC24268.1; GSPDB:GN00019; CBS

A/Experimental source: strain Bristol N2; clone C32F10

C/Genetics: A/Gene: CESP:C32F10.5

A/Map position: 1

A/Intons: 79/3; 292/2; 606/3

C/Superfamily: HMW box homology

Query Match 41.1%; Score 51; DB 2; Length 689;
Best Local Similarity 45.5%; Pred. No. 21;
Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 3 KATYEALQYEAADLAAYKKA 24
Db 616 KARYEAEKKEKKKGGEKAS 637

RESULT 14

S61535 nucleotide-binding head-stalk protein 183K - Giardia lamblia

C/Species: Giardia lamblia
C/Date: 23-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004

C/Accession: S61535

R/Marshall, J.; Holberton, D.V.

J. Cell Sci. 108, 2683-2692, 1995

A/Title: Giardia gene predicts a 183 kDa nucleotide-binding head-stalk protein.

A/Reference number: S61535; MUID:9606038; PMID:7593309

A/Accession: S61535

A/Status: preliminary; nucleic acid sequence not shown

A/Molecule type: DNA

A/Residues: 1-1620 <MAR>

A/Cross-references: UNIPROT:Q24984; EMBL:X79815; NID:9871047; PID:9871048

Query Match 40.3%; Score 50; DB 2; Length 1620;
Best Local Similarity 52.4%; Pred. No. 72;
Matches 11; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Qy 2 AKATYEALQYEAADLAAYK 22
Db 1101 AKQAYEAQIASLTGDLAAKK 1121

RESULT 15

A45981

peptidyl/lysoyl isomerase (EC 5.2.1.8) CYP-40 - human

N/Alternate names: cyclophilin-40; cyclosporin A-binding protein

C/Species: Homo sapiens (man)
C/Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C/Accession: A45981; S36372; S33658

R/Kieffer, L.J.; Seng, T.W.; Li, W.; Osterman, D.G.; Handschumacher, R.E.; Bayney, R.M.

J. Biol. Chem. 268, 12303-12310, 1993

A/Title: Cyclophilin-40, a protein with homology to the P59 component of the steroid r

A/Reference number: A45981; MUID:93286056; PMID:8509368

A/Accession: A45981

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-370 <KIEI>

A/Cross-references: UNIPROT:O08752; GB:U11667; NID:9348909; PIDN:AAA35731.1; PID:93489.


```

RA Ogier J.A., Scholler M., Leproivre Y., Pini A., Sommer P., Klein J.P.;
RT "Complete nucleotide sequence of the sr gene from Streptococcus mutans
RT OM2.175."
RL FEMS Microbiol. Lett. 68:223-228(1990).
DR EMBL; X53647; CAA37697.1; -.
DR PIR; A60988; A60988.
DR HSSP; P11657; 1JMM.
DR InterPro; IPR004829; Csurface_antigen.
DR InterPro; IPR009578; Strept_SA_rep.
DR Pfam; PF06696; Strept_SA_rep; 7.
DR ProDom; PD153432; Csurface_antigen; 2.
SQ SEQUENCE 1556 AA; 171177 MW; 73A3A33736F53A1E CRC64;

Query Match 100.0%; Score 124; DB 2; Length 1556;
Best Local Similarity 100.0%; Pred. No. 5.9e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAKATYEALKOYEADLAAYKKANAA 26
DB 362 NAKATYEALKOYEADLAAYKKANAA 387

RESULT 3
SPAP_STRMU STANDARD; PRT; 1562 AA.
ID SPAP_STRMU STANDARD; PRT; 1562 AA.
AC P23504;
DT 01-NOV-1991 (Rel. 20, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Cell surface antigen I/II precursor.
GN Name=spaf; OrderedLocNames=SMU.610;
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxId=1309;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC STRAIN=NG5 / Serotype C;
RX MEDLINE=90076473; PubMed=2687020; DOI=10.1016/0014-5793(89)81632-1;
RA Kelly C., Evans P., Bergmeier L., Lee S.F., Prognolske-Fox A.,
RA Harris A.C., Aitken A., Bleiweis A.S., Lehner T.;
RT "Sequence analysis of the cloned streptococcal cell surface antigen
RT I/II."
RL FEMS Lett. 258:127-132(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NG5 / Serotype C;
RX MEDLINE=91207143; PubMed=1982405; DOI=10.1016/0003-9969(90)90128-W;
RA Kelly C., Evans P., Ma J.K.C., Bergmeier L.A., Taylor W., Brady L.U.,
RA Lee S.F., Bleiweis A.S., Lehner T.;
RT "Sequencing and characterization of the 185 kDa cell surface antigen
RT of Streptococcus mutans."
RL Arch. Oral Biol. 35:33S-38S(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186; DOI=10.1073/pnas.172501299;
RA Adic D.J., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,
RA Olan Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,
RA Ferretti J.T.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
CC -!- FUNCTION: Surface protein antigen implicated in dental caries.
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (potential).
CC -!- SIMILARITY: Belongs to the spaf/spsp5/spaa family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17390; CAA35253.1; -.
DR EMBL; AE014905; AAN58348.1; -.
DR PIR; S06839; S06839.
DR HSSP; P11657; 1JMM.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR009578; Strept_SA_rep.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF06696; Strept_SA_rep; 7.
DR ProDom; PD153432; Csurface_antigen; 2.
DR TIGRPFAM; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS05847; GRAM_POS_ANCHORING; 1.
KW Antigen; Cell wall; Complete proteome; Dental caries;
KW Direct protein sequencing; Peptidoglycan-anchor; Repeat; Signal.
KW SIGNAL
FT 1 38
FT CHAIN 1 38
FT CHAIN 1 38
FT PROPEP 1532 1532
FT DOMAIN 60 551
FT DOMAIN 220 465
FT DOMAIN 848 964
FT SITE 1529 1533
FT MOD_RES 1532 1532
FT FT 73 73
FT CONFLICT 79 90
FT FT 179 179
FT CONFLICT 183 183
FT FT 227 227
FT CONFLICT 824 824
FT CONFLICT 878 878
FT CONFLICT 984 984
FT CONFLICT 1010 1010
FT CONFLICT 1069 1069
FT CONFLICT 1120 1120
FT CONFLICT 1201 1201
FT CONFLICT 1241 1241
FT CONFLICT 1307 1307
FT CONFLICT 1323 1323
FT CONFLICT 1371 1371
FT CONFLICT 1406 1410
FT CONFLICT 1416 1416
FT CONFLICT 1429 1429
FT CONFLICT 1494 1494
FT CONFLICT 1512 1512
FT CONFLICT 1527 1527
SQ SEQUENCE 1562 AA; 169971 MW; 298B244E7A95F507 CRC64;

Query Match 100.0%; Score 124; DB 1; Length 1562;
Best Local Similarity 100.0%; Pred. No. 5.9e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAKATYEALKOYEADLAAYKKANAA 26
DB 362 NAKATYEALKOYEADLAAYKKANAA 387

RESULT 4
PAC_STRMU STANDARD; PRT; 1565 AA.
ID PAC_STRMU STANDARD; PRT; 1565 AA.
AC P11657;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Pac protein precursor.
GN Name=spaf;
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.

```



```

OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 39-48.
RC STRAIN=MT8148 / Serotype C;
RX MEDLINE=89343654; PubMed=2761390;
RA Okahashi N., Sasaki C., Yoshikawa M., Hamada S., Koga T.;
RT "Molecular characterization of a surface protein antigen gene from
RT serotype C Streptococcus mutans, implicated in dental caries.";
RL Mol. Microbiol. 3:673-678(1989).
CC -1- FUNCTION: Surface protein antigen implicated in dental caries.
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (Potential).
CC -1- SIMILARITY: Belongs to the spaP/spa5/spaA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14490; CAA32652.1; -.
DR PIR; S04729; S04729.
DR PDB; 1UMM; X-ray; A=463-839.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR005578; Strep_SA_rep.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF06696; Strep_SA_rep; 7.
DR Prodom; PD15343; Centface_antigen; 2.
DR TIGRfam; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
DR 3D-structure; Antigen; Cell wall; Dental caries;
KM Direct protein sequencing; Peptidoglycan-anchor; Repeat; Signal.
FT SIGNL 1 38
FT CHAIN 39 1535
FT PROPEP 1536 1565
FT DOMAIN 157 460
FT SITE 835 988
FT MOD_RES 1532 1536
FT MOD_RES 1535 1535
FT TURN 464 464
FT HELIX 465 490
FT TURN 491 495
FT TURN 497 498
FT STRAND 500 500
FT STRAND 506 506
FT STRAND 508 509
FT TURN 514 515
FT STRAND 517 522
FT STRAND 525 527
FT HELIX 529 537
FT HELIX 539 541
FT HELIX 542 546
FT TURN 547 548
FT STRAND 549 550
FT HELIX 552 554
FT HELIX 557 560
FT HELIX 563 565
FT STRAND 566 567
FT STRAND 572 574
FT TURN 575 576
FT TURN 578 579
FT STRAND 583 584
FT TURN 585 587
FT STRAND 588 588
FT TURN 589 589
FT STRAND 593 593
FT TURN 601 602
FT STRAND 604 610
FT STRAND 616 617
FT TURN 618 619

```

```

FT STRAND 620 631
FT TURN 633 634
FT STRAND 642 647
FT HELIX 650 652
FT STRAND 654 657
FT STRAND 664 664
FT STRAND 668 679
FT TURN 680 681
FT STRAND 684 684
FT STRAND 687 694
FT STRAND 699 699
FT STRAND 703 708
FT STRAND 713 715
FT STRAND 717 717
FT TURN 718 719
FT STRAND 722 725
FT TURN 726 727
FT STRAND 728 731
FT TURN 739 740
FT TURN 742 743
FT TURN 754 755
FT TURN 764 765
FT HELIX 766 769
FT STRAND 771 777
FT STRAND 780 788
FT HELIX 789 791
FT TURN 792 792
FT HELIX 795 797
FT TURN 802 803
FT TURN 806 807
FT STRAND 810 810
FT STRAND 813 816
FT STRAND 819 819
FT STRAND 829 829
SQ SEQUENCE 1565 AA; 170781 MW; 4C3B05C809D0C32A CRC64;

Query Match 100.0%; Score 124; DB 1; Length 1565;
Best Local Similarity 100.0%; Pred. No. 5.9e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NAKATYEALKQYEADLAIVKKANNA 26
Db 361 NAKATYEALKQYEADLAIVKKANNA 386

RESULT 5
054183 PRELIMINARY; PRT; 1569 AA.
AC 054183;
DT 01-NOV-1996 (TRENBLREL). 01, Created)
DT 01-NOV-1996 (TRENBLREL). 01, Last sequence update)
DT 01-MAR-2004 (TRENBLREL). 26, Last annotation update)
DE SSpA (Fragment).
GN Name=SSpA,
OS Streptococcus gordonii.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1302;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MS;
RX MEDLINE=96310377; PubMed=8733238;
RA Demuth D.R., Duan Y., Brooke W., Holmes A.R., McNab R.,
RA Jenkinson H.F.;
RT "Tandem genes encode cell surface polypeptides SSpA and SSpB that
RT mediate adhesion of the oral bacterium Streptococcus gordonii to human
RT and bacterial receptors." (1996).
RL Mol. Microbiol. 20:403-413(1996).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL; U40025; AAC44099.1; -.
DR HSSP; P11657; 1UMM.
DR GO; GO:0009986; C:cell surface; IEA.

```

```

DR GO:0005618; C:cell wall; IEA.
DR InterPro: IPR004829; Csurfactant antigen.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR009578; Strep_SA_rep.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR Pfam: PF06696; Strep_SA_rep; 7.
DR ProDom: PD153432; Csurfactant antigen; 2.
DR TIGRfams: TIGR01167; LPXTG_anchor; 1.
DR PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; peptidoglycan-anchor.
FT NON_TER 1
SQ SEQUENCE 1569 AA; 171560 MW; 77F3B189412601C2 CRC64;

Query Match 88.7%; Score 110; DB 2; Length 1569;
Best Local Similarity 100.0%; Pred. No. 4.7e-06;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 AKATYEALKOYEADLAAYKKAN 24
DB 350 AKATYEALKOYEADLAAYKKAN 372

RESULT 6
O54186 PRELIMINARY; PRT; 1499 AA.
ID O54186;
AC O54186;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Streptococcal surface protein B precursor.
GN Name=sspb;
OS Streptococcus gordonii.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_Taxid=1302;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DL1;
RX MEDLINE=96310377; PubMed=8733238;
RA Demuth D.R., Duan Y., Brooks W., Holmes A.R., McNab R.,
  Jenkinson H.F.;
RT "Tandem genes encode cell-surface polypeptides Sspa and Sspb which
  mediate adhesion of the oral bacterium Streptococcus gordonii to human
  and bacterial receptors."
RL Mol. Microbiol. 20:403-413 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=DL1;
RX MEDLINE=20368582; PubMed=1096265; DOI=10.1006/mpat.2000.0369;
RA El-Sabaeny A., Demuth D.R., Park Y., Lamont R.J.;
RT "Environmental conditions modulate the expression of the sspa and sspb
  genes in Streptococcus gordonii."
RL Microb. Pathog. 29:101-113 (2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=DL1;
RX MEDLINE=22983095; PubMed=14617182;
RA Heddle C., Nobbs A.H., Jakubovics N.S., Gal M., Mansell J.P.,
  Dymock D., Jenkinson H.F.;
RT "Host collagen signal induces antigen I/II adhesin and invasion gene
  expression in oral Streptococcus gordonii."
RL Mol. Microbiol. 50:597-607 (2003).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=DL1;
RA Kojander P.E.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=DL1;
RA Jenkinson H.F., Nobbs A.H., Heddle C., Dymock D., Du L.D.,
  Kojander P.E.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RA Jenkinson H.F., Nobbs A.H., Heddle C., Dymock D., Du L.D.,
  Kojander P.E.;

```

```

RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=DL1;
RA Kojander P.E.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -1 SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
  an amide bond (By similarity).
DR EMBL: U40027; AAC44102.3; -.
DR PIR: S70994; S70994.
DR HSPB; P11657; 13MM.
DR GO:0000986; C:cell surface; IEA.
DR GO:0005618; C:cell wall; IEA.
DR InterPro: IPR004829; Csurfactant antigen.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR009578; Strep_SA_rep.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR Pfam: PF06696; Strep_SA_rep; 7.
DR ProDom: PD153432; Csurfactant antigen; 3.
DR TIGRfams: TIGR01167; LPXTG_anchor; 1.
DR PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; peptidoglycan-anchor; Signal.
FT SIGNAL 1 37 Potential.
SQ SEQUENCE 1499 AA; 163973 MW; 6FPB5FADDB2F4FA CRC64;

Query Match 86.3%; Score 107; DB 2; Length 1499;
Best Local Similarity 91.7%; Pred. No. 1.1e-05;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 AKATYEALKOYEADLAAYKKANA 25
DB 361 AKATYEALKOYEADLAAYKKANA 384

RESULT 7
ID SSPS_STRGN STANDARD; PRT; 1500 AA.
AC P16952; O54184;
DT 01-AUG-1990 (Rel. 15, Created)
DT 15-JUN-1998 (Rel. 36, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Agglutinin receptor precursor (SSP-5).
GN Name=ssps; Synonyms=sspb;
OS Streptococcus gordonii.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_Taxid=1302;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M5;
RX MEDLINE=90236997; PubMed=2185241;
RA Demuth D.R., Golub E.E., Malamud D.;
RT "Streptococcal-host interactions. Structural and functional analysis
  of a Streptococcus sanguis receptor for a human salivary
  glycoprotein."
RL J. Biol. Chem. 265:7120-7126 (1990).
RN [2]
RP SEQUENCE FROM N.A., AND REVISIONS.
RC STRAIN=M5;
RX MEDLINE=96310377; PubMed=8733238;
RA Demuth D.R., Duan Y., Brooks W., Holmes A.R., McNab R.,
  Jenkinson H.F.;
RT "Random genes encode cell-surface polypeptides Sspa and Sspb which
  mediate adhesion of the oral bacterium Streptococcus gordonii to human
  and bacterial receptors."
RL Mol. Microbiol. 20:403-413 (1996).
CC -1- FUNCTION: May bind salivary acid residues of salivary agglutinin
  (SAG) in a calcium-dependent reaction. The interaction of SAG with
  its receptor in various oral streptococci modulate bacterial
  colonization of oral tissue and is associated with reduced levels
  of dental caries.
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by

```

```

CC      an amide bond (potential).
CC      -1- SIMILARITY: Belongs to the spaf/spsp/spaa family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U40026; AAC44100.1; -.
DR      PIR; A35186; A35186.
DR      HSSP; P11657; 1JMM.
DR      InterPro; IPR001899; Gram_pos_anchor.
DR      InterPro; IPR009578; Strep_SA_rep.
DR      Pfam; PF00746; Gram_pos_anchor; 1.
DR      Pfam; PF06696; Strep_SA_rep; 7.
DR      ProDom; PD153432; Gsurface_antigen; 2.
DR      TIGRfams; TIGR01167; LPXTG_anchor; 1.
DR      PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW      Calcium-binding; Cell wall; Peptidoglycan-anchor; Repeat; Signal.
FT      SIGNAL 1 38
FT      CHAIN 39 1470 Agglutinin receptor.
FT      PROPEP 1471 1500 Removed by sortase (potential).
FT      DOMAIN 164 470 4 X approximate tandem repeats. HRI.
FT      REPEAT 164 241 1.
FT      REPEAT 242 323 2.
FT      REPEAT 324 405 3.
FT      REPEAT 406 470 4.
FT      DOMAIN 771 887 3 X approximate tandem repeats. PRL.
FT      DOMAIN 1414 1436 Pro-rich (PR2).
FT      CA_BIND 220 235 Potential.
FT      CA_BIND 301 316 Potential.
FT      CA_BIND 931 950 Potential.
FT      CA_BIND 1300 1315 Potential.
FT      SITE 1467 1471 LPXTG sorting signal (potential).
FT      MOD_RES 1470 1470 Pentaglycyl murein peptidoglycan amidated
      threonine (potential).
SQ      SEQUENCE 1500 AA; 164552 MW; DCF190E7D44D889F CRC64;

Query Match 86.3%; Score 107; DB 1; Length 1500;
Best Local Similarity 91.3%; Pred. No. 1.1e-05;
Matches 21; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 AKATYEALKOYEADLAAYKKAN 24
      |||||:|||||:|||||
DB      280 AKATYEALMKOYENDLAAYKKAN 302

RESULT 8
OS4185 PRELIMINARY; PRT; 1575 AA.
ID OS4185;
AC OS4185;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Streptococcal surface protein A precursor.
GN Name=spA;
OS Streptococcus gordonii.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1302;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D1;
RX MEDLINE=96310377; PubMed=8733238;
RA Demuth D.R., Duan Y., Brooks W., Holmes A.R., McNab R.,
RA Jenkinson H.F.;
RT "Tandem genes encode cell-surface polypeptides spA and spB which
RT mediate adhesion of the oral bacterium Streptococcus gordonii to human
RT and bacterial receptors.";
RL Mol. Microbiol. 20:403-413 (1996).

```

```

RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=D1;
RX      MEDLINE=20368582; PubMed=10906265; DOI=10.1006/impat.2000.0369;
RA      El-Sabaeny A., Demuth D.R., Park Y., Lamont R.J.;
RT      "Environmental conditions modulate the expression of the spA and spB
RT      genes in Streptococcus gordonii.";
RL      Microb. Pathog. 29:101-113 (2000).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=D1;
RX      MEDLINE=22983095; PubMed=14617182;
RA      Hedde C., Nobbs A.H., Jakubovics N.S., Gal M., Mansell J.P.,
RA      Dymock D., Jenkinson H.F.;
RT      "Host collagen signal induces antigen I/II adhesin and invasive gene
RT      expression in oral Streptococcus gordonii.";
RL      Mol. Microbiol. 50:597-607 (2003).
RN      [4]
RP      SEQUENCE FROM N.A.
RC      STRAIN=D1;
RA      Jenkinson H.F., Nobbs A.H., Hedde C., Dymock D., Du L.D.,
RA      Kolenbrander P.E.;
RL      Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN      [5]
RP      SEQUENCE FROM N.A.
RC      STRAIN=D1;
RA      Jenkinson H.F., Nobbs A.H., Hedde C., Dymock D., Du L.D.,
RA      Kolenbrander P.E.;
RL      Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN      [6]
RP      SEQUENCE FROM N.A.
RC      STRAIN=D1;
RA      Jenkinson H.F., Nobbs A.H., Hedde C., Dymock D., Du L.D.,
RA      Kolenbrander P.E.;
RL      Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC      -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
      an amide bond (By similarity).
DR      EMBL; U40027; AAC44101.3; -.
DR      PIR; S70995; S70995.
DR      HSSP; P11657; 1JMM.
DR      GO; GO:0009986; C:cell surface; IEA.
DR      GO; GO:0005618; C:cell wall; IEA.
DR      InterPro; IPR004829; Gsurface_antigen.
DR      InterPro; IPR001899; Gram_pos_anchor.
DR      InterPro; IPR009578; Strep_SA_rep.
DR      Pfam; PF00746; Gram_pos_anchor; 1.
DR      Pfam; PF06696; Strep_SA_rep; 7.
DR      ProDom; PD153432; Gsurface_antigen; 2.
DR      TIGRfams; TIGR01167; LPXTG_anchor; 1.
DR      PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW      Cell wall; Peptidoglycan-anchor; Signal.
FT      SIGNAL 1 38 Potential.
SQ      SEQUENCE 1575 AA; 171770 MW; 1C8BA700D3BA2464 CRC64;

Query Match 85.5%; Score 106; DB 2; Length 1575;
Best Local Similarity 95.7%; Pred. No. 1.6e-05;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AKATYEALKOYEADLAAYKKAN 24
      |||||:|||||:|||||
DB      362 AKATYEALKOYEADLAAYKKAN 384

RESULT 9
OS4183 PRELIMINARY; PRT; 1653 AA.
ID OS4183;
AC OS4183;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PAAa.
GN Name=paaH;
OS Streptococcus criceti.

```

```

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1333;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21167676; PubMed=11267788;
RA Tamura H., Kikuchi T., Shirato R., Kato H.;
RT "Cloning and DNA sequencing of the surface protein antigen I/II (Paa)
of Streptococcus citreus."
RL FEMS Microbiol. Lett. 196;251-256(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=E49;
RC PubMed=15329493;
RA Tamura H., Yamada A., Saito H., Murali S., Kato H.;
RT "Identification of another surface protein antigen I/II gene, paaB,
and a putative transcriptional regulator gene, par, from Streptococcus
citreus."
RL Genes Genet. Syst. 79;129-137(2004).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (By similarity).
DR EMBL: AB042239; BAA95000.1; -.
DR HSSP: P11657; 1JMM.
DR GO: GO:0009986; C:cell surface; IEA.
DR InterPro: IPR004829; Csurface_antigen.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR009578; Strept_SA_rep.
DR Pfam: PF00746; Gram_pos_anchor.1.
DR Pfam: PF06696; Strept_SA_rep.11.
DR ProDom: PD153432; Csurface_antigen; 2.
DR TIGRfam: TIGR01167; LpATG_anchor.1.
DR PROSITE: PS50847; GRAM_POS_ANCHORING.1.
DR Cell wall, peptidoglycan-anchor.
SQ SEQUENCE 1653 AA; 180862 MW; 93F28E215BF598DE CRC64;

Query Match 71.8%; Score 89; DB 2; Length 1653;
Best Local Similarity 76.0%; Pred. No. 0.0035;
Matches 19; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 AKATYEALKOYEADLAAYKKNAA 26
DB 530 AKADYEALAOYKDLAAYKKNAA 554

```

```

RL J. Bacteriol. 172;3988-4001(1990).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (Potential).
CC -1- MISCELLANEOUS: Immunodominant determinants are located in the C-
terminal two-thirds of the spa protein.
CC -1- SIMILARITY: Belongs to the spa/sep5/spa family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL: X57841; CAA40973.1; -.
DR EMBL: M38210; AAA26977.1; -.
DR HSSP: P11657; 1JMM.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR009578; Strept_SA_rep.
DR Pfam: PF00746; Gram_pos_anchor.1.
DR Pfam: PF06696; Strept_SA_rep.6.
DR ProDom: PD153432; Csurface_antigen; 1.
DR PROSITE: PS50847; GRAM_POS_ANCHORING.1.
DR Antigen; Cell wall, peptidoglycan-anchor; Repeat; Signal.
FT SIGNAL 1 50
FT CHAIN 51
FT CHAIN ? 1506
FT PROPEP 1507 1528
FT SITE 1503 1507
FT MOD_RES 1506 1506
FT CONFLICT 427 427
FT CONFLICT 431 431
FT CONFLICT 434 434
FT CONFLICT 531 531
FT CONFLICT 600 600
SQ SEQUENCE 1528 AA; 165572 MW; B235F9CD92838ED CRC64;

Query Match 63.7%; Score 79; DB 1; Length 1528;
Best Local Similarity 65.4%; Pred. No. 0.073;
Matches 17; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 NAKATYEALKOYEADLAAYKKNAA 26
DB 292 SAKADYEALAOYKDLAAYKKNAA 317

```

DR InterPro: IPR004829; Surface antigen.
 DR InterPro: IPR001899; Gram_pos_anchor.
 DR InterPro: IPR009578; Strep_SA_rep.
 DR Pfam: PF00746; Gram_pos_anchor; 1.
 DR Pfam: PF06966; Strep_SA_rep; 6.
 DR ProDom: PD153432; Surface antigen; 2.
 DR TIGRfams: TIGR01167; LPXTG_anchor; 1.
 DR PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
 KW Cell wall; Peptidoglycan-anchor.
 SQ SEQUENCE 1566 AA; 170671 MW; 16A14E11BFC85968 CRC64;

Query Match 62.1%; Score 77; DB 2; Length 1566;
 Best Local Similarity 64.0%; Pred. No. 0.14;
 Matches 16; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 AKATYEALKQYEADLAAYKKNAA 26
 DB 365 AKATYEALKQYKQDLAAQSGNAA 389

RESULT 12
 ID Q55308 PRELIMINARY; PRT; 1566 AA.
 AC Q55308;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Surface protein antigen precursor.
 GN Name=pas;
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NC NCBI_TaxID=1310;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MT3791;
 RX MEDLINE=89290018; PubMed=261267; DOI=10.1016/0014-5793(89)80664-7;
 RA Takahashi I., Okahashi N., Sasakawa C., Yoshikawa M., Hamada S.,
 RA Koga T.;
 RT "Homology between surface protein antigen genes of Streptococcus
 RT sobrinus and Streptococcus mutans."
 RL FEBS Lett. 249:383-388(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MT3791;
 RX MEDLINE=91348886; PubMed=1840575;
 RA Tokuda M., Okahashi N., Takahashi I., Nakai M., Nagaoka S.,
 RA Kawagoe M., Koga T.;
 RT "Complete nucleotide sequence of the gene for a surface protein
 RT antigen of Streptococcus sobrinus."
 RL Infect. Immun. 59:3309-3312(1991).
 CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
 CC an amide bond (By similarity).
 CC EMBL: D90354; BAA14368.1; -
 DR PIR: A43607; A43607.

DR HSSP: P11657; IJMM.
 DR GO: GO:0009986; C:cell surface; IEA.
 DR InterPro: IPR004829; Surface antigen.
 DR InterPro: IPR001899; Gram_pos_anchor.
 DR InterPro: IPR009578; Strep_SA_rep.
 DR Pfam: PF00746; Gram_pos_anchor; 1.
 DR Pfam: PF06966; Strep_SA_rep; 6.
 DR ProDom: PD153432; Surface antigen; 2.
 DR TIGRfams: TIGR01167; LPXTG_anchor; 1.
 DR PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
 KW Cell wall; Peptidoglycan-anchor; Signal.
 FT CHAIN 1 38 Potential.
 FT SIGNAL 1 38 Potential.
 SQ SEQUENCE 1566 AA; 170210 MW; F54677C342501C8 CRC64;

Query Match 62.1%; Score 77; DB 2; Length 1566;
 Best Local Similarity 69.6%; Pred. No. 0.14;
 Matches 16; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 AKATYEALKQYEADLAAYKKN 24
 DB 201 AKADYEALKQYKQDLAAVQOAN 223

RESULT 13
 ID Q9KWS1 PRELIMINARY; PRT; 1310 AA.
 AC Q9KWS1;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Pas.
 GN Name=pas;
 OS Streptococcus intermedius.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NC NCBI_TaxID=1338;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CAI-1157;
 RX MEDLINE=21167676; PubMed=1126778;
 RA Tamura H., Kikuchi T., Shirato R., Kato H.;
 RT "Cloning and DNA sequencing of the surface protein antigen I/II (Paa)
 RT of Streptococcus cricetus."
 RL FEMS Microbiol. Lett. 196:251-256(2001).
 CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
 CC an amide bond (By similarity).
 DR EMBL: AB045140; BAA96878.2; -
 DR PIR: C60339; C60339.

DR HSSP: P11657; IJMM.
 DR GO: GO:0009986; C:cell surface; IEA.
 DR GO: GO:0005618; C:cell wall; IEA.
 DR InterPro: IPR004829; Surface antigen.
 DR InterPro: IPR001899; Gram_pos_anchor.
 DR InterPro: IPR009578; Strep_SA_rep.
 DR Pfam: PF00746; Gram_pos_anchor; 1.
 DR Pfam: PF06966; Strep_SA_rep; 3.
 DR ProDom: PD153432; Surface antigen; 1.
 DR TIGRfams: TIGR01167; LPXTG_anchor; 1.
 DR PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
 KW Cell wall; Peptidoglycan-anchor.
 SQ SEQUENCE 1310 AA; 143089 MW; 0021781F5DDB793 CRC64;

Query Match 53.2%; Score 66; DB 2; Length 1310;
 Best Local Similarity 54.2%; Pred. No. 3.6;
 Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 AKATYEALKQYEADLAAYKKN 25
 DB 198 SRAAYETKLAQYQAEIKRVOEANA 221

RESULT 14
 ID Q8SWM4 PRELIMINARY; PRT; 1710 AA.
 AC Q8SWM4; Q9VK26;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE LD28380P (CG16972-PA).
 GN ORFNames=CG16972;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RA Stapleton M., Brooksstein P., Hong L., Abhayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Faltan D., Frise E.,

RA George R., Gonzalez M., Guatin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Murgall C.J., Nunoo J., Pacle J., Pargás V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Cealinker S.,
RA Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10711132; DOI=10.1126/science.287.5461.2185;
RX Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabot G.L.,
RA Abell J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Bresson K.Y., Buesh P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butts K.C., Busam D.A., Butler H., Cadenhead E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke Z., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Davis A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Roeder C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Idegam C.,
RA Jostali M., Kalush F., Karen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Milshina N.V., Mobery C., Norris J., Wostrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Murry D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Petkan G.S., Pan S., Pollard J., Puti V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodruff, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of *Drosophila melanogaster*.";
RT Science 287:2185-2195(2000).
RN [3]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=2242605; PubMed=12537568;
RX Celinker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA George R.A., Adams M., Champe M., Dugan S.P., Frise E., Hogson A.,
RA Patel S., Hoskins R.A., Lavelly T., Murry D.M., Nelson C.R.,
RA Pacle J.M., Pak S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Swinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*
RT *melanogaster* euchromatic genome sequence.";
RT Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [4]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RX Katiner J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celinker S.E.,
RT "The transposable elements of the *Drosophila melanogaster* euchromatin
RT a genomics perspective.";
RT Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [5]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=22426059; PubMed=12537572;
RX Miura S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Katiner J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,

Best Local Similarity 52.0%; Pred. No. 1.3e+02;
Matches 13; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 1 NAKATYEALKOYEADLAAYKKANA 25
Db 1632 NAOXSLEKRLKKA EKDLAAKAAASA 1656

Search completed: October 14, 2005, 14:46:52
Job time : 178 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2005, 14:35:07 ; Search time 22 Seconds
(without alignments)
88.222 Million cell updates/sec

Title: US-10-714-852-1

Perfect score: 124

Sequence: 1 MAKATYERALKQYEADLAAYKANAA 26

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	100.0	1561	3	US-08-894-017-23
2	124	100.0	1561	4	US-09-456-474-23
3	124	100.0	1565	6	5352450-2
4	124	100.0	1565	6	5352450-2
5	124	100.0	1565	6	5352450-2
6	45.2	45.2	17	3	US-08-687-956A-23
7	54	43.5	982	4	US-09-248-588-50
8	54	43.5	982	4	US-09-688-188B-90
9	53	42.7	248	4	US-09-281-417D-90
10	51	41.1	1637	3	US-08-134-000C-5840
11	51	41.1	1637	3	US-09-718-692-2
12	51	41.1	1637	3	US-09-718-852-2
13	49.5	39.9	370	4	US-09-538-092-1286
14	49.5	39.9	370	4	US-09-538-092-1286
15	49.5	39.9	370	4	US-09-949-016-6581
16	49.5	39.9	370	4	US-09-949-016-6581
17	49.5	39.9	370	4	US-09-949-016-6581
18	49.5	39.9	370	4	US-09-949-016-6581
19	48.5	38.7	1043	4	US-09-949-016-6581
20	48.5	38.7	1043	4	US-09-949-016-6581
21	46.5	37.5	111	4	US-09-270-767-46693
22	46.5	37.5	111	4	US-09-270-767-46693
23	46.5	37.5	111	4	US-09-270-767-46693
24	46.5	37.5	111	4	US-09-270-767-46693
25	46.5	37.5	111	4	US-09-270-767-46693
26	46.5	37.5	111	4	US-09-270-767-46693
27	46.5	37.5	111	4	US-09-270-767-46693

28	45.5	36.7	107	4	US-09-270-767-44576	Sequence 44576, A
29	45.5	36.7	160	4	US-08-529-055-33	Sequence 33, Appl
30	45.5	36.7	272	4	US-09-252-991A-26064	Sequence 26064, A
31	45.5	36.7	8991	4	US-08-714-741-32	Sequence 32, Appl
32	45	36.3	160	4	US-09-540-236-2842	Sequence 2842, Ap
33	45	36.3	364	4	US-09-489-039A-12968	Sequence 12968, A
34	45	36.3	429	4	US-09-543-681A-6228	Sequence 6228, Ap
35	44.5	35.9	361	4	US-09-543-681A-5390	Sequence 5390, Ap
36	44.5	35.9	513	4	US-09-724-519-6	Sequence 6, Appl
37	44.5	35.9	513	4	US-09-592-037-6	Sequence 6, Appl
38	44.5	35.9	513	4	US-09-428-156B-6	Sequence 6, Appl
39	44.5	35.9	575	4	US-09-724-519-8	Sequence 8, Appl
40	44.5	35.9	575	4	US-09-592-037-8	Sequence 8, Appl
41	44.5	35.9	575	4	US-09-428-156B-8	Sequence 8, Appl
42	44.5	35.9	1053	4	US-09-724-519-2	Sequence 2, Appl
43	44.5	35.9	1053	4	US-09-592-037-2	Sequence 2, Appl
44	44.5	35.9	1056	4	US-09-595-684B-29	Sequence 29, Appl
45	44.5	35.9	1057	3	US-09-541-782-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-894-017-23
Sequence 23, Application US/08894017
Patent No. 6024958
GENERAL INFORMATION:
APPLICANT: Lehner, Thomas
TITLE OF INVENTION: POLYPEPTIDE FRAGMENTS CAPABLE
OF COMPETITION WITH STREPTOCOCCUS MUTANS ANTIGEN I/II
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESS: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Avenue, NW, suite 5500
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/894,017
FILING DATE: 31-JUL-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00207
FILING DATE: 31-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 25150-20067.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-887-1500
TELEFAX: 202-822-0168
TELEX:
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-894-017-23

Query Match 100.0%; Score 124; DB 3; Length 1561;
Best Local Similarity 100.0%; Pred. No. 1.3e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NAKATYEAALKOYEADLAAYKKANAA 26
Db 361 NAKATYEAALKOYEADLAAYKKANAA 386

RESULT 2

US-09-456-474-23
; Sequence 23, Application US/09456474
; Patent No. 6500433
; GENERAL INFORMATION:
; APPLICANT: Lehner, Thomas
; APPLICANT: Kelly, Charles
; TITLE OF INVENTION: POLYPEPTIDE FRAGMENTS CAPABLE OF
; TITLE OF INVENTION: COMPETITION WITH STREPTOCOCCUS MUTANS ANTIGEN I/II
; FILE REFERENCE: 25150-20067.10
; CURRENT APPLICATION NUMBER: US/09/456.474
; CURRENT FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: US 08/894,017
; PRIOR FILING DATE: 1997-10-20
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 1561
; TYPE: PR1
; ORGANISM: Streptococcus mutans
US-09-456-474-23

Query Match 100.0%; Score 124; DB 4; Length 1561;
Best Local Similarity 100.0%; Pred. No. 1.3e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NAKATYEAALKOYEADLAAYKKANAA 26
Db 361 NAKATYEAALKOYEADLAAYKKANAA 386

RESULT 3

5352450-2
; Patent No. 5352450
; APPLICANT: KOGA, TOSHIHIKO, OKAHASHI, NOBUO, TAKAHASHI, ICHIRO;
; SHIBUYA, KOJI, OHTA, HIROTAKA
; TITLE OF INVENTION: METHOD FOR PREPARING VACCINE FOR DENTAL
; CARRIES AND VACCINAL COMPOSITIONS FOR DENTAL CARRIES USED AS NASAL
; DROP
; NUMBER OF SEQUENCES: 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/529.602
; FILING DATE: 29-MAY-1990
; SEQ ID NO: 2
; LENGTH: 1565
5352450-2

Query Match 100.0%; Score 124; DB 6; Length 1565;
Best Local Similarity 100.0%; Pred. No. 1.3e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NAKATYEAALKOYEADLAAYKKANAA 26
Db 362 NAKATYEAALKOYEADLAAYKKANAA 387

RESULT 4

5352450-2
; Patent No. 5352450
; APPLICANT: KOGA, TOSHIHIKO, OKAHASHI, NOBUO, TAKAHASHI, ICHIRO;
; SHIBUYA, KOJI, OHTA, HIROTAKA
; TITLE OF INVENTION: METHOD FOR PREPARING VACCINE FOR DENTAL
; CARRIES AND VACCINAL COMPOSITIONS FOR DENTAL CARRIES USED AS NASAL
; DROP
; NUMBER OF SEQUENCES: 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/529.602
; FILING DATE: 29-MAY-1990

; SEQ ID NO: 2
; LENGTH: 1565
5352450-2

Query Match 100.0%; Score 124; DB 6; Length 1565;
Best Local Similarity 100.0%; Pred. No. 1.3e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NAKATYEAALKOYEADLAAYKKANAA 26
Db 362 NAKATYEAALKOYEADLAAYKKANAA 387

RESULT 5

US-08-687-956A-23
; Sequence 23, Application US/08687956A
; Patent No. 5861157
; GENERAL INFORMATION:
; APPLICANT: BURNIE, JAMES P
; APPLICANT: MATTHEWS, RUTH C
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: INFECTIONS DUE TO STREPTOCOCCI AND ENTEROCOCCI
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY, MADISON & SUTRO, LLP
; STREET: 1100 NEW YORK AVENUE, N.W., EAST TOWER, 9TH
; STREET: FLOOR
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,956A
; FILING DATE: 29-JUL-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9401689.6
; FILING DATE: 28-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 50865/222892
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/861-3000
; TELEFAX: 202/822-0944
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1566 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus sobrinus
US-08-687-956A-23

Query Match 62.1%; Score 77; DB 2; Length 1566;
Best Local Similarity 64.0%; Pred. No. 0.0028; 7; Indels 0; Gaps 0;
Matches 16; Conservative 2; Mismatches

Oy 2 AKATYEAALKOYEADLAAYKKANAA 26
Db 365 AKATYEAALKOYEADLAAYKKANAA 389

```

RESULT 6
US-09-248-588-50
; Sequence 50, Application US/09248588
; Patent No. 6231864
; GENERAL INFORMATION:
; APPLICANT: Birkett, Ashley J.
; TITLE OF INVENTION: Strategically Modified Hepatitis B Core Proteins and
; FILE REFERENCE: SYN-101 4564/69529
; CURRENT APPLICATION NUMBER: US/09/248,588
; PRIOR FILING DATE: 1999-02-11
; EARLIER APPLICATION NUMBER: 60/074537
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Streptococcus sobrinus
; PUBLICATION INFORMATION:
; JOURNAL: Arch. Oral Biol.
; VOLUME: 35
; PAGES: Suppl. 475-
; DATE: (1990)
US-09-248-588-50

Query Match          45.2%; Score 56; DB 3; Length 17;
Best Local Similarity 75.0%; Pred. No. 0.022;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      2 AKATYEALKQYEADL 17
Db      1 AKADYEALQAYEKDL 16

RESULT 7
US-09-688-188B-90
; Sequence 90, Application US/09688188B
; Patent No. 6656716
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0328
; CURRENT APPLICATION NUMBER: US/09/688,188B
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/291,417
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 60/081,784
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 90
; LENGTH: 982
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-688-188B-90

Query Match          43.5%; Score 54; DB 4; Length 982;
Best Local Similarity 52.4%; Pred. No. 6.3;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy      3 KATYEALKQYEADLAAYKKA 23
Db      651 RSTYENMKVKADLNRYKDA 671

RESULT 8
US-09-291-417D-90
; Sequence 90, Application US/09291417D
; Patent No. 6680170
; GENERAL INFORMATION:

```

```

; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0329
; CURRENT APPLICATION NUMBER: US/09/291,417D
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/081,784
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 90
; LENGTH: 982
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-291-417D-90

Query Match          43.5%; Score 54; DB 4; Length 982;
Best Local Similarity 52.4%; Pred. No. 6.3;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy      3 KATYEALKQYEADLAAYKKA 23
Db      651 RSTYENMKVKADLNRYKDA 671

RESULT 9
US-09-134-000C-5840
; Sequence 5840, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5840
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5840

Query Match          42.7%; Score 53; DB 4; Length 248;
Best Local Similarity 44.7%; Pred. No. 1.7;
Matches 17; Conservative 2; Mismatches 7; Indels 12; Gaps 2;

Qy      1 NAKATYEAA-----LKQYEADLAAY---KKAANA 26
Db      24 NAQACYEAAAGHVVLGVLSKNYETDEAIDMKKYQA 61

RESULT 10
US-09-718-692-2
; Sequence 2, Application US/09718692
; Patent No. 6383796
; GENERAL INFORMATION:
; APPLICANT: Berard, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6383796e1 motor proteins and methods for
; FILE REFERENCE: 1052
; CURRENT APPLICATION NUMBER: US/09/718,692
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1637
; TYPE: PRT

```

```
; ORGANISM: Human
; US-09-718-692-2
Query Match      41.1%; Score 51; DB 3; Length 1637;
Best Local Similarity 41.7%; Pred. No. 35;
Matches 10; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Qy      3 KATYEALIKOYEADLAAYKANA 26
Db      738 OSRYERELKQIAEVAEMKAKVA 761

; RESULT 11
; US-09-718-852-2
; Sequence 2, Application US/09718852
; Patent No. 6426193
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6426193el motor proteins and methods for
; FILE REFERENCE: 1052
; CURRENT APPLICATION NUMBER: US/09/718,852
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1637
; TYPE: PRT
; ORGANISM: Human
; US-09-718-852-2
Query Match      41.1%; Score 51; DB 4; Length 1637;
Best Local Similarity 41.7%; Pred. No. 35;
Matches 10; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Qy      3 KATYEALIKOYEADLAAYKANA 26
Db      738 OSRYERELKQIAEVAEMKAKVA 761

; RESULT 12
; US-09-718-815-2
; Sequence 2, Application US/09718815
; Patent No. 6455293
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6455293el motor proteins and methods for
; FILE REFERENCE: 1052
; CURRENT APPLICATION NUMBER: US/09/718,815
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1637
; TYPE: PRT
; ORGANISM: Human
; US-09-718-815-2
Query Match      41.1%; Score 51; DB 4; Length 1637;
Best Local Similarity 41.7%; Pred. No. 35;
Matches 10; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Qy      3 KATYEALIKOYEADLAAYKANA 26
Db      738 OSRYERELKQIAEVAEMKAKVA 761

; RESULT 13
; US-09-538-092-1286
; Sequence 1286, Application US/09538092
```

```
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSegFormer Version 0.9
; SEQ ID NO 1286
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number Q08752
; US-09-538-092-1286
Query Match      39.9%; Score 49.5; DB 4; Length 370;
Best Local Similarity 46.4%; Pred. No. 9.8;
Matches 13; Conservative 3; Mismatches 7; Indels 5; Gaps 1;

Qy      1 NAKATYEAA-----LKQYEADLAAYKKA 23
Db      306 NTKALYRRAQGQGLKEVDQALADLKKA 333

; RESULT 14
; US-09-949-016-6581
; Sequence 6581, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6581
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-6581
Query Match      39.9%; Score 49.5; DB 4; Length 370;
Best Local Similarity 46.4%; Pred. No. 9.8;
Matches 13; Conservative 3; Mismatches 7; Indels 5; Gaps 1;

Qy      1 NAKATYEAA-----LKQYEADLAAYKKA 23
Db      306 NTKALYRRAQGQGLKEVDQALADLKKA 333

; RESULT 15
; US-09-949-016-11385
; Sequence 11385, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;; FILE REFERENCE: CL001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; CURRENT FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: Fast-Seq for Windows Version 4.0
;; SEQ ID NO: 11385
;; LENGTH: 376
;; TYPE: PRT
;; ORGANISM: Human
US-09-949-016-11385

Query Match 39.9%; Score 49.5; DB 4; Length 376;
Best Local Similarity 46.4%; Pred. No. 9.9;
Matches 13; Conservative 3; Mismatches 7; Indels 5; Gaps 1;

QY 1 NAKATYEA-----LKQYEDLAAVXKA 23
DB 312 NTKALYRRRAQGGWGLKEYDQALADLXKA 339

Search completed: October 14, 2005, 14:42:38
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2005, 14:40:23 ; Search time 64 Seconds
(without alignments)
169.292 Million cell updates/sec

Title: US-10-714-852-1
Perfect score: 124
Sequence: 1 NAKATYBAALKOYEADLAAYKANA 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_PUBCOMB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
21: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	124	100.0	26	US-10-714-852-1	Sequence 1, Appl1
2	124	100.0	1561	US-10-282-122A-72279	Sequence 72279, A
3	57	46.0	1296	US-10-263-929-131	Sequence 131, App
4	57	46.0	1296	US-11-097-143-38835	Sequence 38835, A
5	56	45.2	16	US-09-930-915A-44	Sequence 44, Appl
6	56	45.2	16	US-10-082-014-68	Sequence 68, Appl
7	56	45.2	16	US-10-372-076-69	Sequence 69, Appl
8	56	45.2	16	US-10-806-006-44	Sequence 44, Appl
9	56	45.2	16	US-10-677-074-69	Sequence 69, Appl
10	56	45.2	16	US-10-805-913-44	Sequence 44, Appl
11	56	45.2	16	US-10-732-862A-82	Sequence 82, Appl

	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
	56	56	56	56	56	56	56	55	55	54	54	54	52	52	51	50	50	50	49.5	49.5	49.5	49.5	49.5	49.5	49.5	49.5	49.5	49.5	49.5	49.5	49.5	49.5	49.5	49.5
	45.2	45.2	45.2	45.2	45.2	45.2	45.2	44.4	44.4	43.5	43.5	43.5	41.9	41.9	41.9	40.3	40.3	40.3	39.9	39.9	39.9	39.9	39.9	39.9	39.9	39.9	39.9	39.9	39.9	39.9	39.9	39.9	39.9	39.9
	17	17	17	17	17	17	17	16	16	16	16	16	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
	US-09-930-915A-62	US-10-082-014-143	US-10-372-076-173	US-10-806-006-62	US-10-677-074-173	US-10-805-913-62	US-10-732-862A-193	US-10-42A-599-164152	US-10-369-493-3209	US-09-291-417-90	US-10-725-329-90	US-10-282-122A-60818	US-11-097-143-11352	US-10-282-122A-47540	US-10-915-029-10	US-10-732-923-6764	US-10-041-856-8	US-11-073-203-8	US-10-262-525-8	US-10-511-098-11	US-11-019-829-24	US-10-732-923-6764	US-10-282-122A-6183	US-10-282-122A-60543	US-10-042-865-189	US-09-877-843-94	US-10-282-122A-72386	US-11-097-143-14760	US-10-282-122A-54591	US-10-425-115-267347	US-10-425-115-267348	US-10-425-115-267349	US-10-425-115-278337	US-10-425-115-278344
	Sequence 62, Appl	Sequence 143, App	Sequence 173, App	Sequence 62, Appl	Sequence 173, App	Sequence 62, Appl	Sequence 193, App	Sequence 164152, Sequence 3209, Ap	Sequence 90, Appl	Sequence 90, Appl	Sequence 90, Appl	Sequence 60818, A	Sequence 11352, A	Sequence 47540, A	Sequence 10, Appl	Sequence 6764, Ap	Sequence 46183, A	Sequence 60543, A	Sequence 189, App	Sequence 94, Appl	Sequence 72386, A	Sequence 14760, A	Sequence 54591, A	Sequence 267347, Sequence 267348, Sequence 267349, Sequence 278337, Sequence 278344,										

ALIGNMENTS

RESULT 1
US-10-714-852-1
; Sequence 1, Application US/10714852
; Publication No. US20040132071A1
; GENERAL INFORMATION:
; APPLICANT: SENPOKU, HIDEBOBU
; APPLICANT: MASUZAWA, YUMIKO
; TITLE OF INVENTION: METHOD FOR EXAMINING THE CARRIES RISK
; FILE REFERENCE: 245617US
; CURRENT APPLICATION NUMBER: US/10/714,852
; CURRENT FILING DATE: 2003-11-18
; PRIOR APPLICATION NUMBER: JP 2002-352466
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Streptomyces mutans
US-10-714-852-1

Query Match 100.0%; Score 124; DB 16;
Best Local Similarity 100.0%; Pred. No. 5.6e-11;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAKATYBAALKOYEADLAAYKANA 26
DB 1 NAKATYBAALKOYEADLAAYKANA 26

RESULT 2
US-10-282-122A-72279
; Sequence 72279, Application US/10282122A
; Publication No. US20040029129A1

```
;; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72279
; LENGTH: 1561
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (90)..(90)
; OTHER INFORMATION: X-any amino acid
US-10-282-122A-72279

Query Match      100.0%; Score 124, DB 15, Length 1561;
Best Local Similarity 100.0%; Pred. No. 4, 9e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
;; LENGTH: 1296
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-263-929-131

Query Match      46.0%; Score 57; DB 15, Length 1296;
Best Local Similarity 57.1%; Pred. No. 23;
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY      6 YEAAKQYEADLAAYKKANAA 26
DB      481 FEAAFKQYEAVTASLKQKNPA 501

RESULT 5
US-09-930-915A-44
; Sequence 44, Application US/09930915A
; Publication No. US20030138769A1
; GENERAL INFORMATION:
; APPLICANT: Birkelt, Ashley J.
; TITLE OF INVENTION: IMMUNOGENIC HBC CHIMER PARTICLES HAVING ENHANCED
; FILE REFERENCE: 4564/83501 ICC-102.2 PCT
; CURRENT APPLICATION NUMBER: US/09/930,915A
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/226,867
; PRIOR FILING DATE: 2000-08-22

US-11-097-143-38835
; ORGANISM: DROSOPHILA
; TYPE: PRT
; LENGTH: 1296
; SEQ ID NO 38835
; SOFTWARE: FastSeq for Windows Version 4.0
US-11-097-143-38835

Query Match      46.0%; Score 57; DB 20; Length 1296;
Best Local Similarity 57.1%; Pred. No. 23;
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
```

Query Match 45.2%; Score 56; DB 14; Length 16;

Query Match	45.28;	Score 56;	DB 16;	Length 16;
-------------	--------	-----------	--------	------------

Best Local Similarity 75.0%; Pred. No. 0.26;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AKATYEALKQYEADL 17
||| ||| ||| |||
Db 1 AKADYEAKLAQYEKDL 16

RESULT 10

US-10-805-913-44
; Sequence 44, Application US/10805913
; Publication No. US20040156864A1
; GENERAL INFORMATION:
; APPLICANT: Birkett, Ashley J.
; TITLE OF INVENTION: IMMUNOGENIC HBC CHIMER PARTICLES HAVING ENHANCED
; TITLE OF INVENTION: STABILITY
; FILE REFERENCE: 4564/91645 ICC-102.2 DV II
; CURRENT APPLICATION NUMBER: US/10/805,913
; CURRENT FILING DATE: 2004-03-22
; PRIOR APPLICATION NUMBER: 09/930,915
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: PCT/US01/41759
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/226,867
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,843
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 313
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Streptococcus sobrinus
US-10-805-913-44

Query Match 45.2%; Score 56; DB 16; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.26;

Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AKATYEALKQYEADL 17
||| ||| ||| |||
Db 1 AKADYEAKLAQYEKDL 16

RESULT 11

US-10-732-862A-82
; Sequence 82, Application US/10732862A
; Publication No. US20040146524A1
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, BIRKETT J.
; APPLICANT: Lyons, KateIymne J.
; APPLICANT: Jay, Haron J.
; TITLE OF INVENTION: STABILIZED IMMUNOGENIC HBC CHIMER PARTICLES
; FILE REFERENCE: ICC-136.0 (4564-88881)
; CURRENT APPLICATION NUMBER: US/10/732,862A
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US 60/432,123
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 10/274,616
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US 10/080,299
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 10/082,014
; PRIOR FILING DATE: 2002-02-22
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 82
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Streptococcus sobrinus
US-10-732-862A-82

Query Match 45.2%; Score 56; DB 18; Length 16;

Best Local Similarity 75.0%; Pred. No. 0.26;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AKATYEALKQYEADL 17
||| ||| ||| |||
Db 1 AKADYEAKLAQYEKDL 16

RESULT 12

US-09-930-915A-62
; Sequence 62, Application US/09930915A
; Publication No. US20030138769A1
; GENERAL INFORMATION:
; APPLICANT: Birkett, Ashley J.
; TITLE OF INVENTION: IMMUNOGENIC HBC CHIMER PARTICLES HAVING ENHANCED
; TITLE OF INVENTION: STABILITY
; FILE REFERENCE: 4564/83501 ICC-102.2 PCT
; CURRENT APPLICATION NUMBER: US/09/930,915A
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/226,867
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,843
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 313
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Streptococcus sobrinus
US-09-930-915A-62

Query Match 45.2%; Score 56; DB 10; Length 17;
Best Local Similarity 75.0%; Pred. No. 0.28;

Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AKATYEALKQYEADL 17
||| ||| ||| |||
Db 1 AKADYEAKLAQYEKDL 16

RESULT 13

US-10-082-014-143
; Sequence 143, Application US/10082014
; Publication No. US20030185858A1
; GENERAL INFORMATION:
; APPLICANT: Birkett, Ashley J.
; TITLE OF INVENTION: IMMUNOGENIC HBC CHIMER PARTICLES STABILIZED WITH AN N-TERMINAL C
; FILE REFERENCE: ICC-130.0 4564/85124
; CURRENT APPLICATION NUMBER: US/10/082,014
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 09/930,915
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 143
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Streptococcus sobrinus
US-10-082-014-143

Query Match 45.2%; Score 56; DB 14; Length 17;
Best Local Similarity 75.0%; Pred. No. 0.28;

Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AKATYEALKQYEADL 17
||| ||| ||| |||
Db 1 AKADYEAKLAQYEKDL 16

RESULT 14

US-10-372-076-173
; Sequence 173, Application US/10372076
; Publication No. US20030198645A1


```

; GENERAL INFORMATION:
; APPLICANT: Page, Mark
; TITLE OF INVENTION: STABILIZED HBC CHIMER PARTICLES AS THERAPEUTIC VACCINE FOR
; FILE OF INVENTION: CHRONIC HEPATITIS
; FILE REFERENCE: 4564/87179
; CURRENT APPLICATION NUMBER: US/10/372,076
; PRIOR FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: 10/080,299
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 10/082,014
; PRIOR FILING DATE: 2002-02-22
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 173
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Streptococcus sobrinus
US-10-372-076-173
```

```

Query Match          45.2%; Score 56; DB 14; Length 17;
Best Local Similarity 75.0%; Pred. No. 0.28;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      2 AKATYEALKQYEADL 17
+      |||||
Db      1 AKADYEAKLAQYEKDL 16
```

RESULT 15

```

US-10-806-006-62
; Sequence 62, Application US/10806006
; Publication No. US20040152876A1
; GENERAL INFORMATION:
; APPLICANT: Birkett, Ashley J.
; TITLE OF INVENTION: IMMUNOGENIC HBC CHIMER PARTICLES HAVING ENHANCED
; FILE REFERENCE: 4564/91644 ICC-102.2 DV 1
; CURRENT APPLICATION NUMBER: US/10/806,006
; PRIOR FILING DATE: 2004-03-22
; PRIOR APPLICATION NUMBER: 09/930,915
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: PCT/US01/41759
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/226,867
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,843
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 313
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Streptococcus sobrinus
US-10-806-006-62
```

```

Query Match          45.2%; Score 56; DB 16; Length 17;
Best Local Similarity 75.0%; Pred. No. 0.28;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      2 AKATYEALKQYEADL 17
+      |||||
Db      1 AKADYEAKLAQYEKDL 16
```

Search completed: October 14, 2005, 14:43:51
Job time : 68 secs